Sorting Behavior in Matching Environments: An Experimental Investigation

Bingbing Li[1], Stephanie W. Wang[2], Xiaohan Zhong[3*]

[1] Institute of Population and Labor Economics, The Chinese Academy of Social Sciences, Beijing, China

[2] Department of Economics, University of Pittsburgh, Pittsburgh, USA

[3] Department of Economics, School of Economics and Management, Tsinghua University, Beijing, China

*Corresponding author. E-mail: zhongxh@sem.tsinghua.edu.cn

Abstract

Many matching mechanisms induce non-truth-telling behavior. We study sorting behavior where agents put their most preferred achievable matching objective as the first choice in their preference order list. We examine sorting behavior in a lab experiment under the Serial Dictatorship (strategy-proof) and Boston (non-strategy-proof) mechanisms with complete and incomplete information. Sorting behavior is prevalent under all mechanisms, performs as well as truth-telling and better than other non-truth-telling strategies. Promoting sorting behavior can improve social welfare under Boston mechanisms, in particular ex-ante fairness under incomplete information, but not under Serial Dictatorship mechanisms.

Keywords: Sorting behavior, Boston, Serial dictatorship, Incomplete information

JEL classification: C78, C92, D82

1. Introduction

Non-truth-telling strategic behavior is common under various matching mechanisms. Roth and Sotomayor (1990, Chapter 4) prove, in general, "the impossibility of a 'strategy proof' stable mechanism" in two-sided matching mechanisms. In school choice or college admissions, from the student side, non-truth-telling can be an equilibrium under strategy-proof mechanisms, as well as under non-strategy-proof mechanisms like the Boston (BOS) mechanism or under constrained school choice (Ergin and Sonmez, 2006; Haeringer and Klijn, 2009). Non-truth-telling behavior is also widely observed in lab studies (Chen and Sonmez, 2006; Klijn, Pais and Vorsatz, 2013, etc.), and in the field, for example in China's college admissions (Chen and Kesten, 2017).

In this paper we highlight one specific non-truth-telling behavior: the sorting behavior. *Sorting behavior is for a player to put his or her most preferred achievable matching objective as the first choice in the preference order list.* Sorting behavior is well-defined in a matching system where the stable matching is unique[1]: It is for players to put their unique stable-matched (or fair) objective as the first choice.[2]

Sorting behavior and truth-telling behavior are extremes on the spectrum of strategic behavior. Taking the example of college admissions, truth-telling behavior entails listing *all colleges* in the order of student's true preference, without knowing preferences of all other students and priorities of all colleges. On the contrary, sorting behavior requires information on all other student preferences and all school priorities, to derive the stable matching. In reality, students will try their best to *estimate* their fair college.

At first glance, sorting behavior seems too sophisticated for players to figure out, let alone to implement. However, we need to acknowledge that truth-telling can also be cognitively demanding[3]. In China's college admissions, there are over 2,000 colleges, the feasible choice set for each student can be as large as over 1,000 (depending on whether a college admits students in their residential province). Here we've already ignored choosing dozens of majors for each college. Obviously, students can only

---

1 For example, when the school/college priorities are acyclic (Haeringer and Klijn, 2009, Theorem 7.3).

2 Throughout the paper, "stable" and "fair" are equivalent and used interchangeably. In addition, when we talk about "fair school(s)" of student(s), we always assume the fair matching is unique. If there are multiple stable matches, we define sorting behavior as choosing the fairly-matched school under the *student optimal* stable matching.

3 Li (2017) argues that a strategy-proof mechanism may not be obviously strategy-proof (OSP), therefore, still subject to cognitive limitations of agents. Ashlagi and Gonczarowski (2018) and Troyan (2019) argues that stable matching mechanisms are in general not obviously strategy-proof, and require acyclic assumptions about preferences in order to be OSP-implemented. Note that even under acyclic assumption, mechanisms must be implemented in a specific way to be OSP.

"zoom in" on a relatively small *chosen* set of colleges, a way of thinking that is similar to sorting behavior. In addition, sorting behavior sometimes can be simplified significantly, as in China's college admissions: college priorities are solely determined by the total score rankings on the college entrance exam (CEE). Students can figure out achievable colleges (or rule out unachievable ones), by searching for their score rankings after the CEE through publicly-announced score distributions, or by estimating them through multiple mock exams, if they have to submit preference before the CEE.[4]

Sorting behavior is a Nash equilibrium strategy under various mechanisms. Sorting behavior is an equilibrium strategy for women in a men-optimal stable matching mechanism (Roth and Sotomayor, 1990; Theorem 4.15) and under the Boston mechanism (Ergin and Sonmez, 2006; Theorem 1). Haeringer and Klijn (2009) consider constrained school choice, and their results imply that sorting behavior can be a Nash equilibrium under the TTC mechanism. However, sorting behavior and systematic analysis of non-truth-telling behavior have not been a focus of the matching literature which has prioritized designing strategy-proof mechanisms and promoting truth-telling behavior.

Our experiment is designed to compare sorting behavior to truth-telling and other non-truth-telling behavior under truth-telling and non-truth-telling mechanisms, with regard to prevalence, as well as welfare consequences for individuals and the matching system. We are also interested in how information would affect the choice of sorting behaviors. The experimental design is motivated by China's college admissions system. The system is highly centralized, especially compared to the U.S. system. (Kagel and Roth, 2001; Che and Koh, 2016; Hafalir et al., 2018) In the system, students are uniformly ranked and matched to colleges by their national CEE score. Student preferences over colleges can be submitted before the exam, after the exam but before knowing the exam scores, or after knowing the exam scores. The matching mechanisms used are: The Boston mechanism (known as mechanism "without parallel options") and Serial Dictatorship mechanism (known as mechanism "with parallel options"). However, the SD mechanism is only a constrained one: It is not strategy-proof and only less manipulable than the Boston mechanism (Chen and Kesten, 2017).

We experimentally investigate sorting behavior under four matching environments: Boston with complete information (BOS_C for short) and incomplete information (BOS_I), as well Serial Dictatorship with complete information (SD_C) and incomplete information (SD_I). Complete information corresponds to students submitting their

[4] All the provinces now adopt "preference submission after score announcement". The two latest places making transformation are Beijing (in 2015) and Shanghai (in 2017). All other provinces completed the transformation before 2013.

preference ranking after their exams scores are realized and known, while incomplete information corresponds to submission before scores are realized. College priorities over students are all determined uniformly by student scores, the strongest case of acyclic priorities. Sorting behavior is then (theoretically) unique and (empirically) identifiable.[5] It is to list a student's most preferred achievable school as his or her first choice *given his or her available information*.

We find, first, that sorting behavior is prevalent in all four environments. Under the non-strategy-proof Boston mechanism, about two thirds of students play it. Under the strategy-proof SD mechanism, the proportion is much lower, but still around 20 percent. Information makes no difference on the choice of sorting behaviors. Second, sorting behavior leads to a high first choice admission rate (near 90 percent), and makes the players as well off as truth-telling behavior under any mechanism. Third, an increase in sorting behavior and decrease in other non-truth-telling behavior improves social welfare under the Boston mechanism. In particular, it improves ex-ante fairness (where students with higher unobservable abilities are matched with high-quality schools) under the Boston mechanism with incomplete information. However, this is not the case under the SD mechanism.

The paper is organized as follows: in Section II we lay the theoretical foundations, summarize related literature and propose our testable hypotheses. Section III contains the experimental design and measurements. Section IV reports behavioral patterns with a focus on sorting behavior. Section V focuses on how sorting behavior affects individual well-being, while Section VI examine how promoting sorting behavior would affect social welfare. Section VII looks at a broader class of "misreporting" behavior. Section VIII concludes the paper.

2. Background and Hypotheses

2.1 Theoretical Predictions

Preference Submissions after Exams (Complete Information)

We first look the preference submissions after exams environment. School priorities over students, which are uniquely determined by the ranking of (one-dimensional) exam scores, are common knowledge. Suppose student preferences over

---

[5] More precisely, it is unique with regard to the first choice in the preference order list. As we show later, admission by first choice is often the equilibrium outcome.

schools are also common knowledge. We call this environment a *complete information* setup.[6]

In the school choice/college admissions problem, as in other matching problems, we are concerned about *efficiency* and *stability* of the matching outcomes. In the complete information setup, both stability and efficiency are defined over realized scores or their rankings. In particular, stability (or fairness, the equivalent term often used in the school choice literature) requires that no student-school pairs exist such that the student prefers the school than his/her matched school and the school admits at least one student with a lower *score ranking* than this student, or has a vacancy. Pareto efficiency requires that given the realized scores, there does not exist an alternative matching such that no students become worse off and at least one student becomes better off. We sometimes refer to stability (or fairness) and efficiency under compete information as *ex-post* fairness and *ex-post* efficiency.

We first note that because school priorities are homogenous (based solely on exam score rankings), i.e., acyclic in the strongest sense, the stable matching is unique and ex-post efficient (Haeringer and Klijn, 2009; see also Ergin, 2002, Kesten, 2006). Since the stable (or fair) matching is unique, each student has one specific school as his/he *fair school*, i.e., the school matched to him/her under the stable (or fair) matching. Sorting behavior is therefore well-defined.

Definition 1(sorting behavior under complete information). Sorting behavior is for a student to list his or her (ex-post) fair school (i.e. the school he or she matches under the (ex- post) fair matching) as his or her first choice in his or her preference order list.

Under sorting behavior, a student's second choice and beyond can be any sequence of other non-first-choice schools. However, as we will show, those choices may not matter much at all.

Ergin and Sonmez (2006, Theorem 1) prove that under the Boston mechanism, sorting behavior is a Nash equilibrium which implements the stable matching where students are all admitted by their first choice. Haeringer and Klijn (2009, Theorem 5.4) prove that for the TTC, equilibrium outcomes are *independent* of how many schools a student can list (quotas). And Theorem 6.4 states if school priorities are acyclic, constrained (and unconstrained) TTC implement stable matching outcomes under Nash equilibrium. These two theorems imply that sorting behavior is a Nash equilibrium for

[6] In our experimental design, to be comparable with Chen and Sonmez (2007), we relax the full information requirements on student preferences over schools. However, the environment is still called "complete information" due to information completeness on student score rankings/school priorities.

the TTC, and students are admitted by their first choice under the equilibrium. Finally, note that under homogeneous school priorities, the TTC mechanism is mechanically equivalent to the SD mechanism (Abodulkadiroglu and Sonmez, 2003).

Proposition 1: Under the Boston mechanism and SD mechanism with homogeneous school priorities and complete information, sorting behavior forms a Nash equilibrium and implements (ex-post) fair matching, and students are admitted by their first choice.

Consider a mixture of truth-telling behavior with sorting behavior under the SD mechanism. That is, players either play truth-telling or sorting behavior. It is easy to see that such a strategy profile also forms a Nash equilibrium.

Proposition 2: Any strategy profile combining truth-telling and sorting behavior forms a Nash equilibrium and implements an ex-post fair matching outcome under the SD mechanism with complete information.

Preference Submissions before Exams (Incomplete Information)

We assume that students know the students' *score distributions*, although they do not know the realized scores or score rankings when submitting their preference. We need to adapt several concepts for this environment.

First, we define e*x-ante stability/fairness* based on *expected* exam scores. We assume that the expected score of a student measures his or her true ability and school priorities are based on this. In particular, a match is ex-ante stable or fair if there is no student-school pair such that the student prefers another school to his/her matched school and the school admits at least one student with the lower *expected* score than this student, or has a vacancy. Note that the ex-ante fair match is a deterministic match, where all students are matched 100 percent to their ex-ante fair schools. Note that there exists matching outcomes that are probabilistic, for example, truth-telling equilibrium under SD in incomplete information.

Second, e*x-ante efficiency* (or simply *efficiency*) implies that the expected total payoff of all students across all realizations of the score distribution is maximized. Note that cardinal (instead of ordinal) student preferences over schools are indispensable for ex-ante efficiency.

Note that school priorities based on expected score rankings are still homogeneous. Sorting behavior then is still well-defined.

Definition 2 (sorting behavior under incomplete information). Sorting behavior is for a student to list his or her ex-ante fair school (i.e. the school matched under the ex-ante fair matching) as his or her first choice in his or her preference order list.

It is obvious that if all students list their ex-ante fair school as their first choice, the matching outcome must be ex-ante fair, under either the BOS or SD mechanism. However, sorting behavior is usually not a Nash equilibrium. Lien et al. (2017, Proposition 3.3) prove that under the Boston mechanism with incomplete information (BOS_I), if students have homogeneous preferences over schools, and each school has one slot, the Boston mechanism implements ex-ante fair matching outcomes *only if* sorting behavior is used by all students except the one with the lowest score. Furthermore, ex-ante fairness can only be implemented under very restrictive conditions, i.e., almost no competition (i.e., any overlap on realized scores) between students (Theorem 3.2). The implication of these two theorems is that sorting behavior is almost surely not a Nash equilibrium under BOS_I. Under SD_I, since truth-telling is a dominant strategy, although weakly, it is easy to prove that sorting behavior may not be a Nash equilibrium[7].

Proposition 3. Under the Boston and SD mechanisms with incomplete information, sorting behavior (if used) can achieve ex-ante fairness, but it may not be a Nash equilibrium strategy.

Although sorting behavior may not be a Nash equilibrium, it can still serve as a focal point, especially when there are no obvious equilibrium strategies, e.g., under BOS_I.

[7] Consider one player (student A) deviates from sorting behavior to truth-telling, given other players still playing sorting behavior. Suppose for simplicity all students have the same ordinal preference. Suppose student A's realized score can be larger than another student's, student B's, who has a higher expected score and a more preferable fair school. If everyone chooses sorting behavior, all get their ex-ante fair school. If player A deviates from sorting behavior to truth-telling, he can, with a positive probability, enter into the ex-ante fair school belonging to student B, a more preferable school. Then truth-telling must be strictly better than sorting, given that it is weakly dominant. Sorting behavior is not a Nash equilibrium.

2.2 Related Literature

Understanding Non-truth-telling Behavior

Chen and Sonmez (2006) single out several patterns of non-truth-telling behavior under Boston, TTC and Gale-Shapley (GS, also called Deferred Acceptance (DA)) mechanisms, e.g., small school bias (SSB), district school bias (DSB) and the "Minneapolis strategy" (i.e., "make the first choice a true favorite and the other two 'realistic'''), and find that the Minneapolis strategy is better than other non-truth-telling behavior for the players. Pais and Pinter (2008) highlights "priority school bias" (PSB, i.e., students ranking schools where they have priority higher in the submitted rank) under Boston, TTC and GS mechanisms with various information. Calsamiglia et al. (2010) focus on truncated truth-telling, under constrained school choice. Klijn et al. (2013) study how an individual's risk preference influences his/her strategy, especially the choice of protective strategy (i.e., maxmin strategy) under the GS and BOS mechanisms. Those papers usually do not consider welfare effects of non-truth-telling behavior they highlighted (except for Chen and Sonmez (2006)), and none of these focus on sorting behavior as a general pattern of non-truth-telling behavior.

There is a growing body of literature addressing non-truth-telling behavior under truth-telling mechanisms, mostly by exploring field data. Some of them believe that non-truth-telling behavior is usually a mistake and harmful for the players. Rees-Jones (2018) and Rees-Jones and Skowronek (2018) present evidence that students misrepresent their preferences under the medical residency match (around 20%). Hassidim et al. (2018) find obvious misrepresentation (namely dropping and flipping) in the Israeli Psychology Master's Match (IPMM). These studies also explore factors that predict strategic behavior. A common finding, as summarized in Hassidim, et al. (2018), is that students with lower rankings tend to play strategically more often. Chen and Pereyra (2019) define self-selection, resembling our sorting behavior, as not applying to some schools despite them being desirable if a student believes her chances of being assigned to them are zero. Using data from the Mexico City high school match, they find evidence that self-selection exposes students from low socio-economic backgrounds to strategic mistakes.

Artemov et al. (2017) argue that non-truth-telling behavior under truth-telling mechanisms may not be harmful. It finds that in Australian college admissions, a non-negligible fraction of applicants adopt strategies that are non-truth-telling and unambiguously dominated; yet the majority of these 'mistakes' are payoff irrelevant. They define the stable-response strategy (SRS) to be any strategy that ranks one's most preferred feasible college ahead of other feasible colleges. Note that SRS includes truth-telling strategy and sorting behavior, among others. They prove that non-truth-telling

SRS can be part of a robust equilibrium which is also asymptotically stable. Our results are in line with Artemov et al. (2017): sorting behavior can be payoff irrelevant, in the sense that it is as good as truth-telling for the players.

For the Boston mechanism, Pathak and Sonmez (2008) analyze behavior of sincere players who report their true preferences and of sophisticated players who play a best response. Any sophisticated student weakly prefers her assignment under the Pareto-dominant Nash equilibrium of the Boston mechanism to her assignment under the student-optimal stable (i.e., GS) mechanism. They find that sorting behavior (i.e., their equilibrium strategy) is even better than truth-telling under the Boston mechanism with complete information. Featherstone and Niederle (2014) explore the Boston mechanism (as well as others) with incomplete information, and argue that the non-truth-telling equilibrium is difficult to achieve. They focus on bias in the students' second choice - the "skip the middle" bias. However, when looking at first choice as we do, students do well in figuring out their equilibrium strategy (see Table 2). Wu and Zhong (2020) explore the relation between mismatch (i.e., unfair matching) and misreport (i.e., non-sorting behavior) at the individual level in Chinese college admissions using field data, and find that risk attitude, information deficiency, and preferential policy can lead to misreporting which in turn leads to mismatch.

Pan (2019) studies sorting behavior, and finds that it is affected by overconfidence especially under the Boston mechanism with preference submission before exam scores are known (i.e., BOS_I). More overconfident students play more aggressively and are matched to better colleges, while the matching outcome under BOS_I is more ex-ante unfair than under other environments. Pan only considers homogenous ordinal student preferences, while in our design student preferences are only partially aligned. Thus, we find that students do quite well in making their first choice, and there is little evidence of overconfidence.[8] Overconfidence may have a larger impact on matching outcomes when student preferences are homogeneous since they are competing for exactly the same schools.

Matching mechanisms with Incomplete Information

Some of the previous literature has argued that the Boston mechanism with incomplete information (i.e., BOS_I) may be superior in ex-ante efficiency and fairness to truth-telling mechanisms such as DA/SD. Under the Boston mechanism, students can

[8] In Section VII, we measure "misreport" which captures how students' first choice may deviate from their fair schools. Under BOS_I, more than 60 percent of students fairly report their first choice, only 20 percent of students up-report (i.e., putting a school ranked higher than fair school as the first choice), and 17 percent of students even down-report (i.e., putting a school ranked lower than fair school as the first choice).

reveal either their preference intensity or internal ability through manipulating their preference order lists, while a truth-telling mechanism prevents them from doing so.

Edril and Ergin (2008) found that when random tie-breaking is introduced, GS mechanism may generate *inefficient* stable outcomes. Abdulkadiroglu et al. (2011) proved that when students have the same ordinal preferences (but different cardinal preferences) and schools use a random tie-breaking rule, the Boston mechanism Pareto dominates GS mechanism. In China's college admissions, BOS_I can be more efficient than other mechanisms under conditions consistent with Abdulkadiroglu et al. (2011) (Lien et al., 2016; Chen, 2017). Lien et al. (2017) raised the issue of ex-ante fairness, and proved that BOS_I can be more ex-ante fair than other mechanisms, but the difference is small.[9] In lab experiments, Lien et al. (2016) provided evidence for both ex-ante efficiency and fairness advantages of BOS_I, while Pan (2019) found negative results for the ex-ante fairness of BOS_I.

Our experimental environment is a mixture of Abdulkadiroglu et al. (2011), Lien et al. (2017) and Pan (2019): student scores are different ex-ante unlike Abdulkadiroglu et al, (2011), while students have heterogeneous ordinal preferences unlike Lien et al. (2017) and Pan (2019). We are interested whether sorting behavior emerges in this more realistic setup and how such behavior affects individual and social welfare.

2.3 Hypotheses

We now propose hypotheses for our experimental results based on theoretical predictions and previous empirical results in the literature.

Hypothesis 1. Under both complete and incomplete information, there will be: (i) a significant fraction of players who choose sorting behavior; and (ii) more sorting behavior observed under the BOS mechanism than under the SD mechanism.

For complete information, the hypothesis is derived from Proposition 1 and 2. Sorting behavior is the NE strategy under both BOS_C and SD_C. Furthermore, it can be the unique equilibrium under BOS, while under SD, it is an equilibrium strategy weakly dominated by truth-telling. Under incomplete information, while there is no obvious NE strategy under BOS, sorting behavior can be the focal-point strategy. Truth-telling is still the dominant strategy under SD.

---

[9] Rees-Jones (2017) discusses the similar idea under the GS(DA) mechanism. It argues that some strategic misrepresentation of preferences can facilitate positive assortative matching when student quality is imperfectly observed.

Hypothesis 2. (i) Sorting behavior is more likely to achieve the ex-ante/ex-post fair matching schools under incomplete/complete information for those who choose it. (ii) The likelihood is even higher under the BOS than under the SD mechanism.

Hypothesis 3. (i) By promoting sorting behavior (at the margin), i.e., replacing non-sorting behavior by sorting behavior, the matching outcome under either BOS or SD with complete/incomplete information can achieve more ex-post/ex-ante fair matching outcomes for the whole system. (ii) The effect is larger under the BOS mechanism than under the SD mechanism.

These two hypotheses are motivated by Propositions 1, 2 and 3. Proposition 1 and 3 state that if *all* the students choose sorting behavior, then the matching outcome is ex-post or ex-ante fair under any of the four mechanisms. In addition, under the SD mechanism with complete information, if students choose either sorting or truth-telling behavior, the outcome is still ex-post fair for all students. However, if some students choose another behavior, those who choose sorting behavior are not guaranteed their fair schools. For example, it may happen if other students with higher scores (mistakenly) list the same school as their first choice. Therefore, Hypothesis 2 and 3 only holds under some restrictive assumptions. Nevertheless, we are still interested in how different behaviors perform within a mechanism and how the whole system performs when the behavior mixture changes. Finally, given the possible prevalence of truth-telling under SD, we hypothesize that sorting behavior may not change the matching outcomes as much as under BOS.

A final note on ex-ante efficiency. While ex-post (Pareto) efficiency of a matching outcome is compatible with ex-post fairness (under acyclic school priorities), ex-ante efficiency is a stricter concept and may be different from ex-ante fairness. Ex-ante efficiency depends on cardinal student preferences, while ex-ante fairness only takes ordinal preferences into account; ex-ante fairness takes into account the score distributions (at least expected scores) while ex-ante efficiency does not.[10] Our focus is ex-ante/ex-post fairness, but we still report ex-ante efficiency results for completeness.

[10] As an example, consider students all having ex-ante the same score distribution, then any matching outcome is ex-ante fair/stable, but ex-ante efficiency depends on (heterogeneous) cardinal preferences and may correspond to one specific matching, as illustrated in Abdulkadiroglu et al. (2011) and Lien et al. (2016). Now suppose the score distributions of all students are different. Depending on how we shift the score distributions, ex-ante fairness would correspond to different matchings, while the matching corresponding to ex-ante efficiency is irrelevant of score distributions and unchanged.

3. Experimental Design and Measurement

We first describe our experiment design. We then explain how we identify sorting behavior, and describe the method for evaluating matching outcomes.

3.1 Experimental Design

We implement a 2x2 design, either the Boston or SD mechanism under either a complete or incomplete information environment. For the complete information environment, preference submission is done after the exam and all the students' scores and their rankings are known. For the incomplete information environment, preference submission is done before the scores are known. For each treatment, we have 2 or 3 sessions, with 36 students in each session. We have 10 sessions in total and each student only participated in one session.

Information/Mechanism	Boston	Serial Dictatorship
Complete	2 (N=72)	2 (N=72)
Incomplete	3 (N=108)	3 (N=108)

In the complete information setting, the scores for the 36 students in each session are independently drawn without replacement from integers between 105 and 140. Each student is informed of his/her ranking. In the incomplete information environment, the *estimated* scores for the 36 students in each session are independently drawn without replacement from integers between 105 and 140. Each student is informed of his/her *estimated* ranking. The actual score for each student is drawn from a uniform distribution -10 to +10 around his/her estimated score. The actual ranking is not revealed until the preference lists are already submitted.

In each session, there are 36 school slots across 7 schools A-G: 3 slots each at A and B, 5 slots at C and E, 6 slots at D and F, and 8 slots in G. Schools always prefer students with higher (realized) scores. All students prefer A and B to the other five schools. Some students prefer A to B and some prefer B to A, but the proportion is not publicly known. Each student only knows his/her own preference ranking, but not others'. The payoff structure is similar to Chen and Sönmez (2006) and the payoffs obtained are symmetric, i.e., each student gets the same payoff for the same preference ranking. The actual monetary payment in RMB are obtained by multiplying the payoffs in the table by 5. The payoff table is shown below (for more details see Table 3 in appendix).

	Preference Ranking						
Payoff	1	2	3	4	5	6	7
	16	13	11	9	7	5	2

We ran paper-and-pencil experiments at Tsinghua University in China on May 24th (2 sessions), June 1st (4 sessions) and June 2nd (4 sessions) of 2012. Each session lasted approximately an hour with a baseline participation fee of 20 RMB. All sessions were conducted at the Tsinghua University, School of Economics and Managements' Experimental Economics Laboratory (ESPEL).

We also collected information from the students in a post-experiment survey including their experience (origin province, whether or not they took the CEE, and year of college entrance exam if they did, gender, age, and major). We also use the three series of paired lotteries in Tanaka et al. (2010) for an incentivized elicitation of prospect theory parameters.

3.2 Identifying Sorting Behavior

Sorting behavior is for a player to put his most preferred achievable school as the first choice. Note that this definition does not put restrictions on the second, third and any other choices. In our setup, because students do not have full information on other students' preference (even if they know their realized or expected scores), the unique stable-matched (or fair) school is not fully observable for them. We consider two alternative measures of sorting behavior, by relaxing the theoretical definition. First, we assume students only rely on available information to figure out their fair school. In particular, we define all schools a student is *likely* to be admitted to as his or her *achievable school set*. Sorting behavior then chooses the most preferred school in one's achievable school set as the first choice. Therefore, if a student is ranked (ex-ante or ex-post) among the top 6, the total slot number of school A and B, which all students prefer to other schools, their achievable school set is {A, B, C, D, E, F, G}. Otherwise their achievable school set is {C, D, E, F, G}. Among their achievable school set, they choose their most preferred one according to the endowed payoffs. For example, under an incomplete information setup, a student with an estimated score ranking of 5th, 16 as the payoff for B, 13 as the payoff for A, would be regarded as choosing sorting behavior if he chooses school B as the first choice. Or, under a complete information setup, a student with a realized score ranking as 10th with payoffs C=9, D=11, E=7, F=5, G=2 is choosing sorting behavior if he chooses school D as the first choice.

Second, or alternatively, we can identify whether a student put his (ex-ante/ex-post) fair school *in our experimental design* as his first choice. The shortcoming of this *de facto* sorting behavior is obvious: students may not deliberately choose it due to lack of information. Yet it is still useful, because it answers the "what if" question: i.e., how sorting behavior (or deviation from it) would affect individual matching outcome, if the fair school *were* fully observable by the player? In the later parts, we refer to the first measure of sorting behavior still as "sorting behavior", and the second measure as "*fair reporting*". Note that two measures converge when information on student preferences are fully revealed.

Sorting behavior and truth-telling can overlap. It happens when a student's true first choice coincides with his fair school. Truth-telling is a stricter concept in that it requires all schools to be listed in the true preference order. If a behavior can be both sorting behavior or truth-telling, we identify it as truth-telling.

A key issue for identifying sorting behavior concerns identifying ex-post or ex-ante fair matching for each student. Ex-ante fair school can be derived by ranking students according to their estimated scores, and (hypothetically) match them with schools by using the SD mechanism. The matched school for each student under this procedure is his/her ex-ante fair school. Ex-post fair school can be calculated in a similar way, by using the SD mechanism but ranking students according to their realized scores.

3.3 Measuring Matching Outcomes

For our purpose, it is important not only to measure (or evaluate) matching outcomes for the whole system, as is commonly done in the literature, but also at the individual level.

Measuring Individual Matching Outcomes

We consider two measures: one is related to fairness and the other to efficiency.

The *degree of mismatch* is defined as the gap between the preference ranking of a student's ex-post or ex-ante fair school and his/her actually matched school. Note that the sign of the degree of mismatch can be negative or positive, indicating down-matching (i.e., matched with worse-than-fair school) or up-matching (i.e., matched with better-than-fair school) respectively. For example, a student's (ex-post or ex-ante) fair school is C, which in his payoff table is ranked 3th, yet he is matched to school A, which is ranked first in his payoff table. Then his degree of mismatch is 3-1=2, and he is up-matched.

The *relative payoff* is defined as a student's realized payoff divided by his payoff under the fair match. This measure of the "real monetary gain" complements the degree of mismatch.

Measuring Aggregate Matching Outcomes

Efficiency. The (ex-ante) efficiency is calculated as the payoff per capita averaged over all players in a matching system.

Ex-post Fairness. We adopt two measures of ex-post fairness.

(i) *Ex-post Fairness by number of blocking pairs.* In general, a blocking pair is a student-school pair in which the student prefers the school to his own matched school, and the school either has a vacancy or gives a higher priority to this student than another student it admits. Here the school priority is solely determined by the students' realized scores. In our set-up, we can identify student-school pair (i, S) as a blocking pair if student i prefers school S to his matched school and his score is above the *minimum* score of all the students matched to school S. We then count the number of blocking pairs in a matching system and average across all students. The lower the number of blocking pairs, the fairer the matching outcome.

(ii) *Ex-post Fairness by average absolute value of degree of mismatch.* This measure is calculated by averaging the *absolute value* of the degree of mismatch across all students. A higher degree of mismatch implies a lower level of fairness for the matching system.[11]

Ex-ante Fairness. Ex-ante fairness is defined just like ex-post fairness, except that the fair matchings are defined based on expected student scores instead of realized scores. We still use two alternative measures, i.e., average degree of mismatch, and average number of blocking pairs.

Simulation Considerations. Under the BOS and SD with incomplete information environments, students only know the distribution of scores. Yet students know their realized scores, and then submit their preference rankings in the complete information environments. Thus we need to compare their matching outcomes under the same *prior*

[11] Some discussions on relations between those two fairness measures are deserved. Although the two measures share the same spirit, they may not always be consistent. For example, consider 3 students, 1, 2 and 3, who are matched to 3 schools, A, B, and C. All the students have the same preference: A>B>C, and all the schools have the same priority: 1>2>3. It is easy to calculate that among all 6 possible matching outcomes, there are three matching outcomes with the highest average degree of mismatch (i.e., highest unfairness by measure (ii)): (3=A, 2=B, 1=C), (2=A, 3=B, 1=C), and (3=A, 1=B, 2=C), with the value of 4/3. While by measure (i), the number of blocking pairs, only (3=A, 2=B, 1=C) has the highest degree of unfairness, with the value of 1 (versus 2/3 for the other two). One may argue that (3=A, 2=B, 1=C) should be the only one most unfair, because it "reverses the whole matching order". Yet some may argue that (3=A, 2=B, 1=C) can be fairer than (2=A, 3=B, 1=C) and (3=A, 1=B, 2=C): In (3=A, 2=B, 1=C), at least student 2 is matched to his fair school, while in (2=A, 3=B, 1=C) and (3=A, 1=B, 2=C), all students are mismatched.

score realizations. We need to consider the counterfactuals, i.e., how the students (or subjects) would respond to all the possible score realizations they may experience.[12]

One conjecture would be that students only respond to their realized score rankings: students would play the same strategy as the one with the same realized score ranking observed in the lab. However, it would be naïve to assume that the realized score ranking is the only determinant of behavior. We therefore generate five different profiles of student behavior using empirical patterns with regard to students' first and even second choice observed in the data. More details about the simulation methods are in Appendix A.

4. Results: Patterns of Behavior

Table 1 shows the distribution of behavior for each of the four mechanisms (BOS_I/C, SD_I/C). 65% and 69% of all students play truth-telling under SD_I and SD_C. Only 1% plays truth-telling under BOS_I, and the proportion is 11% for BOS_C.

Table 1: Distribution of student behavior

	BOS_I		BOS_C		SD_I		SD_C	
	#	%	#	%	#	%	#	%
All Students								
Truth-telling	1	0.93	8	11.11	70	64.81	50	69.44
Non-truth-telling	107	99.07	64	88.89	38	35.19	22	30.56
Sorting	68	62.96	47	65.28	19	17.59	17	23.61
Wilcoxon Rank-Sum test for proportion of sorting	BOS_C > BOS_I > SD_C > SD_I (p=0.7521) (p=0.0000) (p=0.3240)							
Top 6 Students								
Truth-telling	1	5.56	7	58.33	15	83.33	10	83.33
Non-truth-telling	17	94.44	5	41.67	3	16.67	2	16.67
Sorting	15	83.33	3	25.00	3	16.67	2	16.67
Risk averse	2	11.11	2	16.67	0	0.00	0	0.00
Students Below Top 6								
Truth-telling	0	0.00	1	1.67	55	61.11	40	66.67
Non-truth-telling	90	100.00	59	98.33	35	38.89	20	33.33
Sorting	53	58.89	44	73.33	16	17.78	15	25.00
Risk seeking type-I	11	12.22	1	1.67	6	6.67	2	3.33
Risk seeking type-II	5	5.56	2	3.33	2	2.22	0	0.00

[12] For the mechanisms under incomplete information, we only need to realize the matching outcome by large enough times according to the score distribution, each time using the same strategies of all students we observe (once) in the lab. We *can* also generate behaviors through simulation by abstracting their behavior patterns. We choose to generate behavior patterns for both complete and incomplete information environments.

Safe choice	7	7.78	5	8.33	5	5.56	0	0.00
Equal slot switch	6	6.67	6	10.00	3	3.33	1	1.67
Less slot switch	8	8.89	1	1.67	3	3.33	2	3.33

We now focus on non-truth-telling behavior. Under BOS_I and BOS_C, 63% and 65% of all students play sorting behavior. Under SD_I and SD_C, sorting behavior is still the most frequently chosen non-truth-telling behavior at 18% and 24%. The proportion of sorting behavior is significantly higher under BOS(_I/C) than under SD(_I/C) (with a p-value of 0.0000 under Wilcoxon rank-sum test).

Result 1 (on Hypothesis 1). Sorting behavior is chosen under all four mechanisms (BOS/SD_I/C), with a significant proportion (18%~65%). It is significantly more frequently chosen under BOS(_I/C) (63~65%) than under SD(_I/C) (18%~24%).

We also divide students into two groups: Top 6 or Below Top 6, according to their expected/realized scores under incomplete/complete information. For Top 6 students, truth-telling and sorting are similar: Both involve listing one's true first choice first. However, students tend to choose sorting behavior much more frequently than truth telling under BOS, in particular, BOS_I. That is, students tend to choose their first choice truthfully, but choose other choices in a way inconsistent with their true preference order.

We also identify some other non-truth telling behavior, for Top 6 students and students below top 6 separately. For Top 6 students, the only observed pattern beside truth-telling and sorting behavior is what we called "risk averse" behavior. That is, students put their less preferred one between school A and B as their first choice. Only 11% and 16% students under BOS_I and BOS_C play it, and none play it under SD mechanisms.

For students below top 6, truth-telling is less popular compared with Top 6 students, even under SD mechanism. Some of them switch to sorting behavior, and some are diverted to other non-truth-telling behavior. This is particularly true for BOS_I, where score uncertainties induce (possibly) risk-related behaviors (risk-seeking or risk-avoiding). Nevertheless, sorting behavior still prevails in all four mechanisms and dominates in the two BOS mechanisms.

Other non-truth telling patterns can also be identified. Here, "risk seeking Type-I" is defined as choosing the most preferred school (between A and B) as their first choice. "Risk seeking type-II" is defined as choosing the second preferred school (between A and B) as their first choice. Another pattern we highlight is "safe choice", which means choosing a school which has more slots than what the school sorting behavior would choose. For example, when sorting behavior requires one to choose school C or E, he/she chooses among D, F, G. For remained behaviors, we can categorize them into two: switching to schools with slots equal to or less than their fair school. These two patterns are hard to explain: they are either chosen randomly, or involve some high level thinking. Finally, note that no students put their least preferred school first.

Table 4-5 contains regression analysis to explore factors affecting the emergence of truth-telling and sorting behavior respectively. Table 4 is on determinants of truth-telling. Truth-telling is more prevalent under SD than Boston mechanisms. Within Boston mechanisms, it is more prevalent under complete information than incomplete information. Under complete information, students have a certain estimation of their rankings. Therefore, for students with high rankings (e.g., the Top 6), they tend to use truth-telling more frequently than under incomplete information (see also Table 1). Within SD mechanisms, truth-telling does not differ between two information settings, reflecting that truth-telling is always a dominant strategy. The behavioral parameter σ has significantly positive effects on the choice of truth-telling, implying that more risk averse players (with lower σ) are less likely to be truth-telling.

Table 5 is on determinants of sorting behavior. Contrasting to truth-telling, sorting behavior is more prevalent under Boston mechanisms. Information has no influence on the emergence of sorting behavior, for either Boston or SD mechanism. Sorting behavior was not influenced by any other factors we include in the regression.

5. Results: Welfare Consequences

In this section we compare welfare consequences of truth-telling and sorting behaviors for individuals *within each mechanism*. We start by examining the first choice admission rate, i.e., the proportion of students admitted by their first choice. Sorting behavior is a strategy for players who aim to be admitted by their first choice, as stated in Proposition 1. Table 2 shows whether this is the case. For all students (or all behaviors), the first choice admission rate is higher under BOS_I/C than under SD_I/C, with the proportion around 80% in the former and around 30% in the latter (Column (1)). Truth-telling (Column (2)) do not generate high first choice admission rates under SD mechanisms, understandably. It seems surprising that it does generate high first choice admission under the Boston mechanisms, but note that very few players play this strategy and almost all of them are top students (Table 1). Sorting behavior generates a high first choice admission rate of 85%-94% under every mechanism (Column (3)).

Table 2: Proportions of Top Choice Match in Different Mechanisms (%)

	All Students	Truth-telling	Sorting	Other Non-truth-telling	Fair-reporting

Treatment	(1)	(2)	(3)	(4)	(5)
BOS_I	81.48	57	92.59	62.74	93.56
BOS_C	88.89	87.5	93.62	76.47	97.78
SD_I	35.63	18.58	84.97	49.11	88.78
SD_C	31.94	16	88.24	0.00	100.00

Note: Percent of top choice match for BOS_I and SD_I is the average value after 200 simulations of score distribution.

Figure 1[13] compares the matching outcome, i.e., the mean degree of mismatch, of different strategies. (See also Table 6 for more details). We calculate the degree of mismatch at the individual level induced by truth-telling, sorting behavior, and all other non-truth-telling behavior. A larger (positive) degree of mismatch is more desirable for the players. Sorting behavior performs no worse than truth-telling under any environments, and even better under BOS_I and SD_C (p-value of Wilcoxon rank-sum test is p=0.0001 and p=0.0020). On the contrary, other non-truth-telling strategies are worse than sorting behavior under any environment (Panel B in Table 6). In addition, sorting behavior is also moderately safe, as measured by the variance of degree of mismatch (Panel C in Table 6).

Sorting behavior not only led to higher-ranking schools (or higher payoffs) for students, it also induces fair schools. In Figure 1 and Panel A of Table 6, we see that sorting behavior on average results in *zero* degree of mismatch.

[13] All the confidential intervals (the red lines in the all figures) are at 95% level.

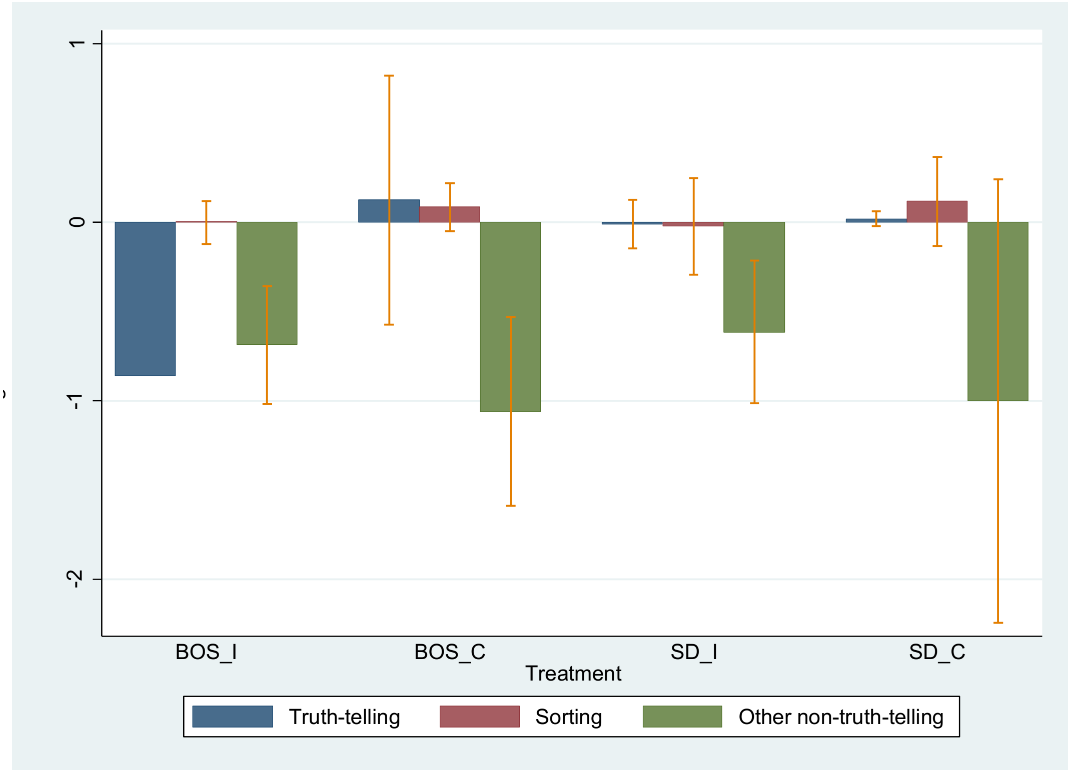
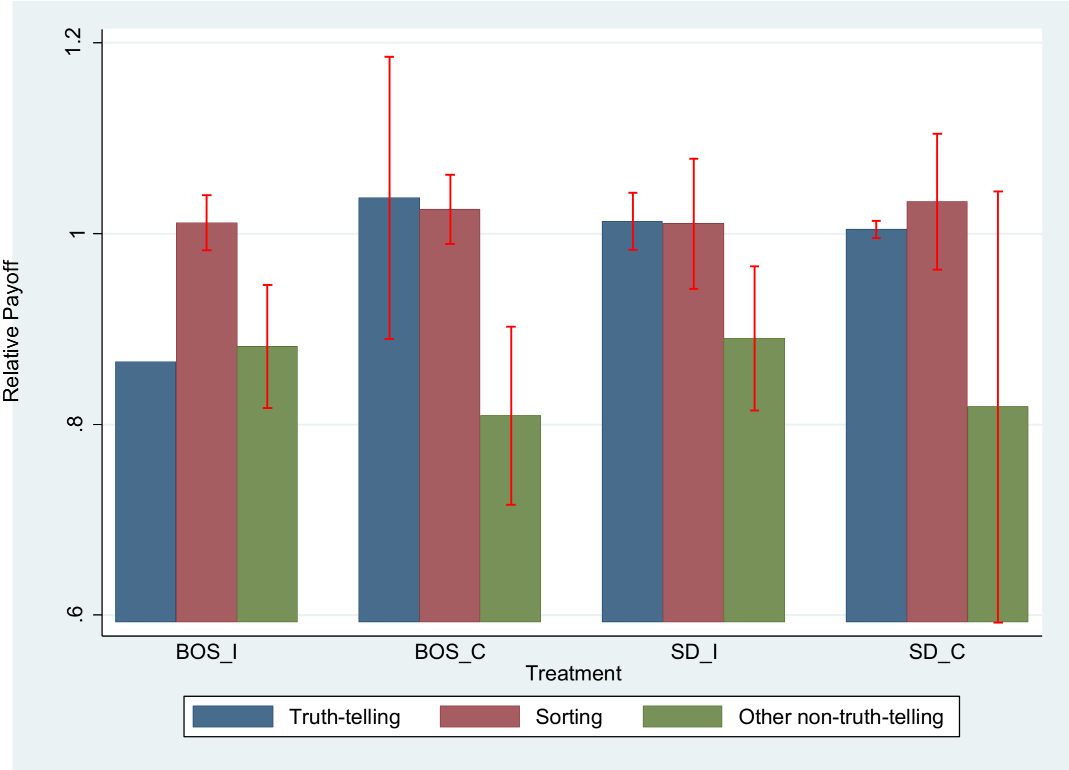


Fig. 1: Degree of Mismatch of Different Strategies under Four Mechanisms

Figure 2 (and Table 7) compare the (ex-ante) efficiency consequences of truth-telling and sorting behavior. The result is the same as for fairness. Sorting behavior performs equally well as truth-telling under BOS_C, SD_I/C, and better than truth-telling under BOS_I. Other non-truth-telling strategies always perform the worst.



Note: Relative payoff is defined as a student's realized payoff divided by his payoff under the fair match.

Fig. 2: Relative Payoffs of Different Strategies

We also consider the welfare consequences of different strategies for Top 6 students and students below top 6 separately (not shown in the table). There are essentially no significant differences among truth-telling, sorting, and other non-truth-telling strategies for Top 6 students. Therefore, the results are driven mainly by students below top 6.

Result 2 (on Hypothesis 2). Sorting behavior induce individual well-being equal to or higher than truth-telling, and higher than other non-truth-telling strategies. In addition, sorting behavior also leads to fair matching.

6. Promoting Sorting Behavior? A Social Welfare View

In Section 5 we showed that sorting behavior performs as well as truth-telling or even better under all mechanisms. Does this imply that policy makers should promote sorting behavior to improve social welfare, in particular fairness? We know, from Proposition 1 and 3, if all students play sorting behavior, ex-ante or ex-post fairness is achieved. However, in reality, it would be hard to persuade all students to choose sorting behavior, due to informational or cognitive constraint. Yet promoting sorting behavior *at the margin* may have ambiguous effects. Sorting behavior may produce negative (or positive) externality on other players. If the negative externality dominates at the margin, promoting sorting behavior may actually decrease social welfare. Furthermore, the effect of promoting sorting behavior may also depend on the mechanisms we use.

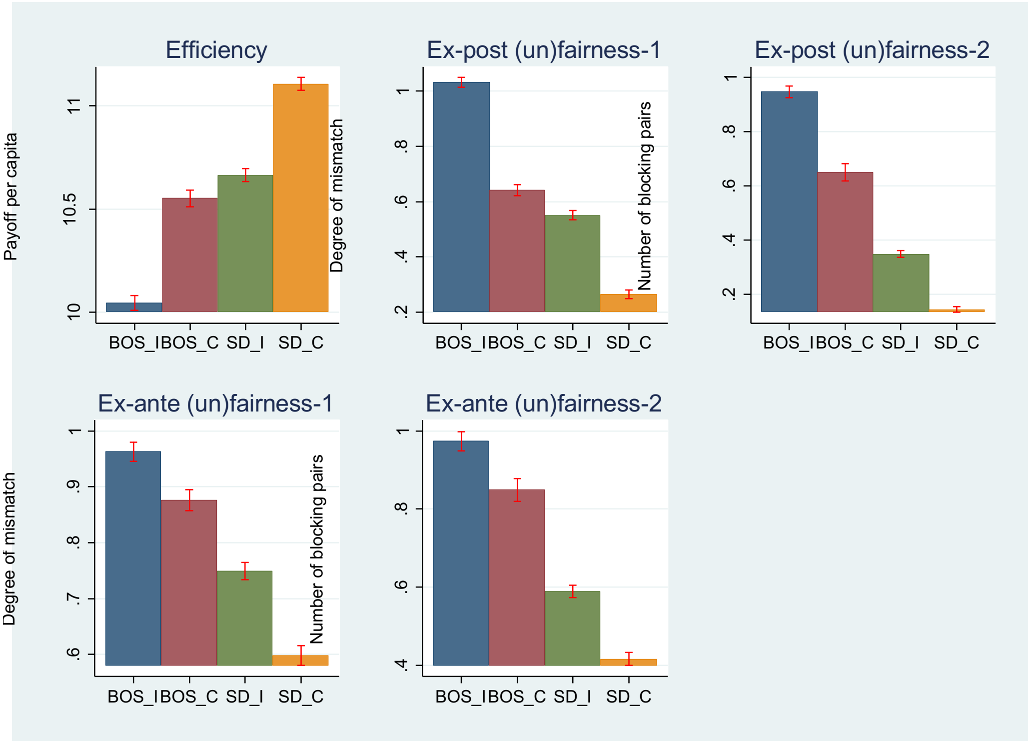
Social Welfare across Environments

As a first look, we examine the social welfare of the original (i.e., without policy interventions) matching outcomes of the four environments. The results are shown in Figure 3 (and Table 8). Simulations are used and scenario 2' (in Appendix A) is considered for the social welfare implications of promoting sorting behaviors (see the next subsection). In all environments, and by all measures (ex-ante efficiency, ex-ante and ex-post fairness), the outcome is consistent: SD_C is always the best, followed by SD_I and BOS_C while BOS_I is the worst.[14] Note that sorting behavior is more frequent under BOS than under SD. Therefore, those results seem to not favor sorting behavior as a path towards social welfare: more sorting behavior leads to more unfair matchings *across mechanisms*.

However, the correlation between sorting behavior frequencies and social welfare of the whole system across mechanisms does not necessarily imply that promoting sorting behavior cannot improve social welfare *for any given mechanism*. In our next subsection, we ask, for any given environment, what happens when we increase the frequency of sorting behavior by replacing truth-telling or other non-truth-telling behavior.

---

[14] In particular, sorting behavior does not help BOS_I to achieve (ex-ante) efficiency or fairness, which contradicts to Lien et al. (2016, 2017) but is aligned with Pan (2019), which suspects the "social value" of sorting behavior.



Note: Efficiency is measure by payoff per capita. Ex-ante or ex-post (un)fairness is measured by degree of mismatch (1) or the number of blocking pairs (2).

Fig. 3: Social Welfare Comparison Among Four Mechanisms

Promoting Sorting Behavior: A Counterfactual Analysis

Since the Boston mechanisms are non-strategy-proof, our counterfactual test only considers the switch between sorting behavior and other non-truth-telling behavior. Yet we also include the all-truth-telling case (i.e., all players play truth-telling) as a benchmark. Under the SD mechanisms, we instead consider the switch between sorting behavior and truth-telling, because other non-truth-telling behavior only plays a minor role - the proportion of other non-truth-telling is much lower under SD than under BOS.[15] We consider increasing or decreasing the proportion of sorting by 6 percent (roughly 2 persons in one session) and 12 percent. We also consider the case of all-sorting behavior.

Figure 4 (and Table 9) shows the results when we increase the proportion of sorting behavior under BOS_I/C. All the welfare measures, including ex-ante efficiency,

[15] Because sorting behavior only characterizes students' first choice, we consider a simulation scenario applicable to this analysis, which is scenario 2' in Appendix A.

ex-ante and ex-post fairness improve significantly (see Table 9), while if we decrease the proportion, the effect is opposite as expected.

For example, under BOS_I, when we increase the proportion of sorting behavior by 6 percent (roughly one tenth of its original proportion), ex-ante efficiency increased by 2 percent (Table 9). Therefore, one percentage increase in sorting behavior (at the original proportion of 63%) increases ex ante efficiency by roughly 0.33 percentage. Ex-ante and ex-post fairness improve more, by about 5-9 percent, i.e., 0.8-1.5 percentage increase for one percentage increase in sorting behavior. When the proportion of sorting behavior increases by 12=2*6 percent, the percentage change is roughly doubled.

The welfare improvement under BOS_C is even larger, especially for ex-ante and ex-post fairness. One percentage increase in sorting behavior increases the ex-ante and ex-post fairness by 1.5-2.5 percent, larger than that under BOS_I. Since BOS_C has already outperformed BOS_I, BOS_C will outperform BOS_I with the same magnitude of small increase in sorting behavior. However, when all the students play sorting behavior, BOS_I outperforms BOS_C in ex-ante efficiency, as we would expect.

We also consider what happens when all the players play truth-telling, or all the students choose sorting behavior. Compared with all-sorting behavior, all truth-telling behavior has higher efficiency, ex-ante and ex-post fairness under BOS_I/C, except for one ex-ante fairness measure under BOS_I.

Note: Efficiency is measure by payoff per capita. Ex-ante or ex-post (un)fairness is measured by degree of mismatch (1) or the number of blocking pairs (2).

Fig. 4: The Effect of Changing the Proportion of Sorting Behavior under Boston Mechanisms

Figure 5 (and Table 10) shows the welfare implications of replacing truth-telling by sorting behavior under SD mechanisms. An increase or decrease in sorting behavior at the margin has almost no effect on social welfare by any measure. For some significant changes, the direction can be non-monotonic. Furthermore, all-sorting behavior performs the worst among all the mixtures in ex-ante and ex-post fairness measured by the number of blocking pairs, while all truth-telling behavior generates the highest social welfare by any measure. As a whole, promoting sorting behavior does not improve social welfare under SD mechanisms.

Note: Efficiency is measured by payoff per capita. Ex-ante or ex-post (un)fairness is measured by degree of mismatch (1) or the number of blocking pairs (2).

Fig. 5: The Effect of Changing the Proportion of Sorting Behavior under SD Mechanisms

We also consider changing the behavior of students below top 6 only. As we have seen, Top 6 students are usually not affected by changing behavior. The welfare consequences are essentially the same as we change behaviors of all students under BOS_I/C. Under SD_I/C, changing the behavior of students below Top 6 to sorting performs better than changing all students' behavior to sorting.

As a whole, promoting sorting behavior helps in reaching desirable matching outcomes under the Boston mechanisms, but has almost no effect under SD mechanisms. Under the Boston mechanisms, it helps both at the margin and at the extreme where all players choose sorting behavior.[16]

Result 3 (on Hypothesis 3). Promoting sorting behavior at the margin improves social welfare (ex-ante efficiency, ex ante and ex post fairness) under BOS(I/C), but has almost no effect under SD(I/C). Hypotheses 3 is (partially) supported.

7. Alternative Measure of Sorting Behavior

In previous sections we focus on one measure of sorting behavior, i.e., choosing the most preferred school within the *set of achievable schools given the available information*. In this section we focus on the other measure, *fair-reporting*, i.e., students choose a school as their first choice that *turns out to be* their fair school.

We can extent the definition of fair-reporting to include *all* non-truth-telling behavior. The *degree of misreport* measures the difference between the preference ranking of the fair school and the first choice. For example, if a student's fair school is a school ranked No 6 in his/her preference list, while his/her first choice is a school ranked No 5, then the degree of misreport is 6-5=1>0. If the degree of mismatch is positive, then the student is said to *up-report*, meaning that he or she lists a school ranked higher than his or her fair school. Otherwise he/she *down-reports*. Therefore, all non-truth-telling behavior is divided into three types: up-reporting, down-reporting, and fair-reporting.[17]

[16] Note that although all truth-telling strategy profile seems the best under any environment, it is not an equilibrium under the Boston mechanism. Therefore, promoting sorting behavior is still a good alternative (if not the only) under those non-strategy-proof mechanisms.

[17] It is not surprising that fair-reporting and sorting behavior (defined in Section 3.2) are highly overlapping. In fact, among all the non-truth-telling behavior, 49.35% is both, 19.9% is either, while other 30.7% is neither.

Figure 6 shows the proportion of various misreporting behavior, as well as truth-telling behavior. Under BOS_C/I, fair-reporting behavior dominates other behavioral patterns, with a proportion of over 60 percent. Under SD_C/I, although truth-telling is dominant, fair-reporting is still prevalent and accounts for nearly 20 percent of all players. There are also significant proportions of up-reporting and down-reporting (10-20 percent).

Table 11 further explores the determinants of fair-reporting, parallel to Table 5 on the determinants of sorting behavior. The result is also similar as Table 5. The Boston mechanisms generate more fair-reporting than SD mechanisms. Within BOS or SD mechanism, information makes no difference.

Table 12 explores the determinants of various misreporting behaviors further. It shows how the *degree* of misreport changes across mechanisms and with other school or student characteristics. To focus on non-truth-telling behavior, truth-telling behavior is excluded from the sample. Misreporting behavior does not differ across four mechanisms. For school characteristics, if a student's fair school has more slots, he or she is more willing to take the risk of up-reporting a higher-ranked school as first choice. For student characteristics, a student with a higher expected/realized score ranking tends to up-report more, and students from Economics and Management school also up-report more.

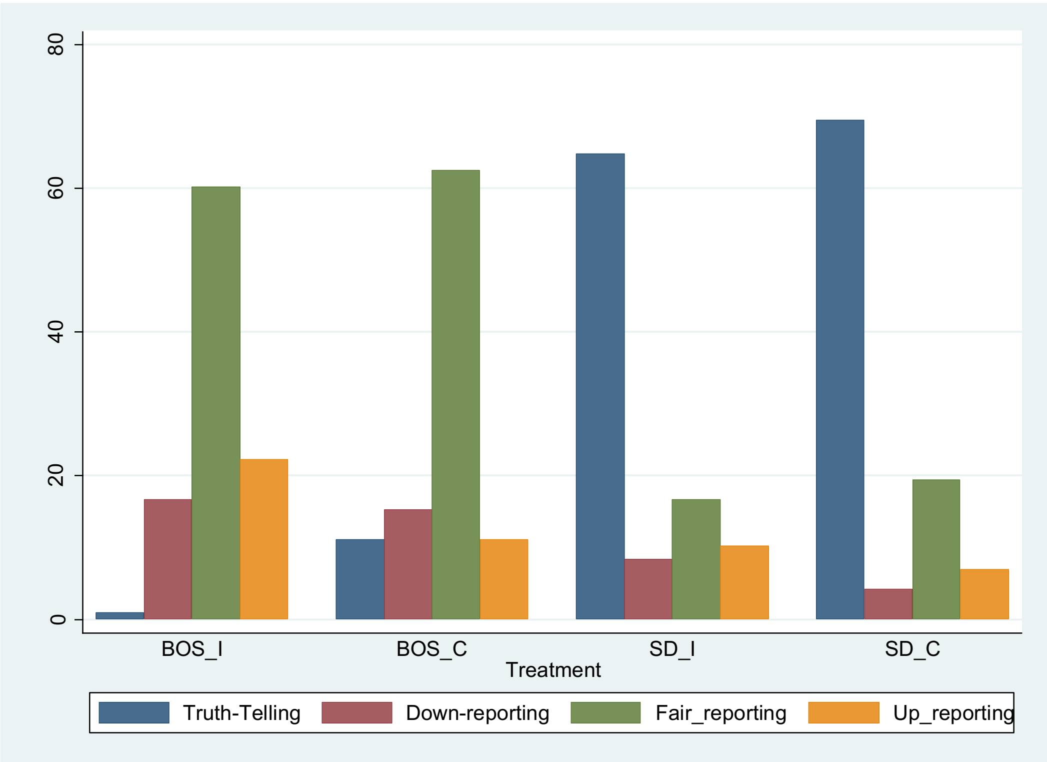


Fig. 6: Proportions of Misreporting and Truth-telling

What happens to the players when he or she fair-reports or mis-reports? Column (5) in Table 2 shows that fair-reporting results in a high first choice admission percentage under all mechanisms. Under SD_C it is 100 percent, while under other mechanisms it is close to or above 90 percent. Figure 7 and Table 13 compare fairness consequences of misreporting under the four mechanisms using the degree of mismatch. Fair-reporting almost always generates fair matching. Up-reporting results in a positive degree of mismatch, or an up-match, while down-reporting results in a negative degree of mismatch, or down-match. Truth-telling also generates fair matching under BOS_C and SD_I/C, but down-matching under BOS_I.

Although up-reporting generates better matching outcome for the player on average, as in Pan (2019), the benefit has its costs: the variance of the degree of mismatch is also higher under up-reporting than under fair-reporting and other strategies (Table 8). Fair-reporting results in lower matching than up-reporting, but it is safer. We also evaluate the effect of changing the proportion of sorting behavior under various mechanisms (as in Figure 4 and 5). The results are similar with our first measure of “sorting behavior” (Figure 8-9).

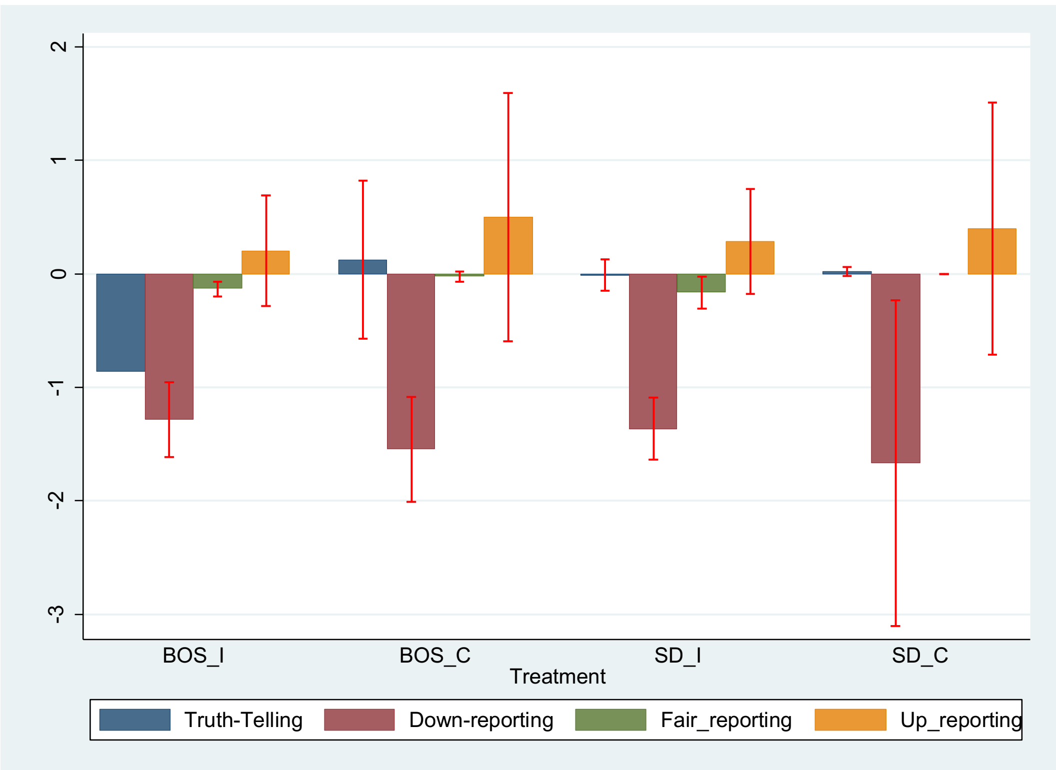


Figure 7: Degree of Mismatch under Misreporting

8. Conclusion

While much of the literature has focused on the conditions for strategy-proof mechanisms and truth-telling behavior, non-strategy-proof mechanisms are still being used and non-truth-telling behavior is prevalent in many matching markets. In this paper we highlight one important non-truth-telling strategy: the sorting behavior, which is the agents listing their most preferred achievable matching objective as their first choice.

Our experiment results verify its frequent use in both non-strategy-proof mechanisms (such as the Boston mechanism) and strategy-proof mechanisms (such as Serial Dictatorship), under complete or incomplete information. By using sorting behavior, ex-ante or ex-post fairness for either individuals or the whole system becomes easier to achieve, and (ex-ante) efficiency may also be improved.

Promoting sorting behavior can be an alternative way to improve social welfare, especially when the mechanism is not strategy-proof. In fact, the Chinese college admissions system is still a non-truth-telling mechanism, despite its transition from the Boston mechanism to the (constrained) SD mechanism. High school teachers, parents, consulting firms, and even colleges have been involved to help students figure out their fair college, by providing information, advice, or even algorithm service. There are other examples for the use of sorting behavior or strategy. In dating service companies, client information is collected and analyzed so that each client is provided with a short list of dating candidates. A PhD graduate on the academic job market gets advice from the advisor to target "ideal" colleges.

So how can sorting behavior be promoted? Information provision is important and the government can play a role. In China's college admissions system, the transition from preference submissions before exams to preference submissions after exams, led by local governments, supposedly helps students to submit their preference order lists "more precisely". Wu and Zhong (2014) find that female students benefit from this transition, perhaps because they are more risk averse than male and the transition helped them to fill the gap. Poor students, especially in rural areas, may have information disadvantages. The government can provide free information services, probably through training programs for parents or school teachers, to help students get their "ideal" colleges (See for example, Ding, et al., 2021).

References

Abdulkadiroğlu, A., Che, Y. K., & Yasuda, Y. (2011). Resolving conflicting preferences in school choice: The" Boston mechanism" reconsidered. *American Economic Review*, *101*(1), 399-410.

Abdulkadiroğlu, A., & Sönmez, T. (2003). School choice: A mechanism design approach. *American economic review*, *93*(3), 729-747.

Ashlagi, I., & Gonczarowski, Y. A. (2018). Stable matching mechanisms are not obviously strategy-proof. *Journal of Economic Theory*, *177*, 405-425.

Calsamiglia, C., Haeringer, G., & Klijn, F. (2010). Constrained school choice: An experimental study. *American Economic Review*, *100*(4), 1860-1874.

Che, Y. K., & Koh, Y. (2016). Decentralized college admissions. *Journal of Political Economy*, *124*(5), 1295-1338.

Chen, L., & Pereyra, J. S. (2019). Self-selection in school choice. *Games and Economic Behavior*, *117*, 59-81.

Chen, Y., & Kesten, O. (2017). Chinese college admissions and school choice reforms: A theoretical analysis. *Journal of Political Economy*, *125*(1), 99-139.

Chen, Y., & Sönmez, T. (2006). School choice: an experimental study. *Journal of Economic theory*, *127*(1), 202-231.

Ding, Y., Du, L., Li, W., Wu., Y., Yang, J., & Ye., X. (2021). The Impact of Informational Interventions on College –Major Choice: Evidence from a Large-Scale Randomized Experiment. *China Economic Quarterly*, 21(6), 2239-2262. (in Chinese)

Erdil, A., & Ergin, H. (2008). What's the matter with tie-breaking? Improving efficiency in school choice. *American Economic Review*, *98*(3), 669-689.

Ergin, H. I. (2002). Efficient resource allocation on the basis of priorities. *Econometrica*, *70*(6), 2489-2497.

Ergin, H., & Sönmez, T. (2006). Games of school choice under the Boston mechanism. *Journal of public Economics*, *90*(1-2), 215-237.

Haeringer, G., & Klijn, F. (2009). Constrained school choice. *Journal of Economic theory*, *144*(5), 1921-1947.

Hafalir, I. E., Hakimov, R., Kübler, D., & Kurino, M. (2018). College admissions with entrance exams: Centralized versus decentralized. *Journal of Economic Theory*, *176*, 886-934.

Hassidim, A., Marciano, D., Romm, A., & Shorrer, R. I. (2017). The mechanism is truthful, why aren't you?. *American Economic Review*, *107*(5), 220-224.

Kagel, J. H., & Roth, A. E. (2000). The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment. *The Quarterly Journal of Economics*, *115*(1), 201-235.

Kesten, O. (2006). On two competing mechanisms for priority-based allocation problems. *Journal of Economic Theory*, *127*(1), 155-171.

Klijn, F., Pais, J., & Vorsatz, M. (2013). Preference intensities and risk aversion in school choice: A laboratory experiment. *Experimental Economics*, *16*(1), 1-22.

Li, S. (2017). Obviously strategy-proof mechanisms. *American Economic Review*, *107*(11), 3257-3287.

Lien, J. W., Zheng, J., & Zhong, X. (2016). Preference submission timing in school choice matching: testing fairness and efficiency in the laboratory. *Experimental Economics*, *19*(1), 116-150.

Lien, J. W., Zheng, J., & Zhong, X. (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.

Pais, J., & Pintér, Á. (2008). School choice and information: An experimental study on matching mechanisms. *Games and Economic Behavior*, *64*(1), 303-328.

Pan, S. (2019). The instability of matching with overconfident agents. *Games and Economic Behavior*, *113*, 396-415.

Rees-Jones, A. (2017). Mistaken play in the deferred acceptance algorithm: Implications for positive assortative matching. *American Economic Review*, *107*(5), 225-229.

Rees-Jones, A. (2018). Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match. *Games and Economic Behavior*, *108*, 317-330.

Rees-Jones, A., & Skowronek, S. (2018). An experimental investigation of preference misrepresentation in the residency match. *Proceedings of the National Academy of Sciences*, *115*(45), 11471-11476.

Roth, A.E., & Sotomayor, M.A.O. (1990). *Two Sided Matching: A Study in Game-Theoretic Modelling and Analysis*. New York: Cambridge University Press.

Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American economic review*, *100*(1), 557-571.

Troyan, P. (2019). Obviously Strategy‐Proof Implementation Of Top Trading Cycles. *International Economic Review*, *60*(3), 1249-1261.

Wu, B., & Zhong, X. (2014). Institutions and Gender Difference in China’s College Admission. *Journal of Tsinghua University (Philosophy and Social Sciences)*, 29(132), 140-156. (in Chinese)

Wu, B., & Zhong, X. (2020). Matching inequality and strategic behavior under the Boston mechanism: Evidence from China's college admissions. *Games and Economic Behavior*, *123*, 1-21.

Appendix A: Simulation Methods for Evaluating Matching Outcomes

We first describe our simulation process, and then we will describe scenarios we use to simulate preference submission behavior.

The simulation steps are as follows.

Step1. Randomly draw all the students' realized score rankings from the given score distribution.

Step2. Simulate students' preference submission based on expected/realized scores under incomplete/complete information through each of the four scenarios we will describe below.

Step3. Match according to the realized scores and simulated preference submission behavior of all students under the Boston or SD mechanism.

Step4. Simulate step 1-3 200 times.

Scenario 1

For the incomplete information treatments, the matching is done according to each realized score ranking. Student preference submission behavior is fixed to what we observe in the data.

For the complete information treatments, we assume a student's behavior is only determined by his or her realized score ranking, and randomly choose from one of the two sessions of each treatment the observed preference submission behavior (for all seven choices) for the same score ranking.

Scenario 2

We use the distribution of truth-telling and up/down/fair reporting strategies for the first choice for simulating student choices. In particular, for each treatment, we calculate the proportion of truth-telling and the distribution of non-truth-telling reporting strategies by looking at the degree of misreport (w.r.t. expected/realized scores for incomplete/complete-information treatments). We then assign students to be truth-telling according to the proportion of truth-telling for each treatment. For non-truth-telling students, we assign students to up/fair/down-reporting the first choice according to the distributions of non-truth-telling reporting strategies observed in the data. We randomly assign the 2nd-7th choices for each non-truth-telling student.

Scenario 2'

We use the distribution of truth-telling, sorting behavior and other behavior patterns revealed in Table 1. In particular, for each treatment, we calculate the proportion of truth-telling, and the distribution of sorting and other non-truth-telling strategies by looking at students' first choice. We distinguish top 6 students and students below top 6. We then assign students to a specific behavioral pattern according to the proportion of that pattern of behavior for each treatment observed in the data (i.e., Table 1). We randomly assign the 2nd-7th choices for each non-truth-telling student.

Scenario 3

Scenario 3 also simulates the non-truth-telling students' second choice after simulating the first choice as in Scenario 2. We assign second choices to the students to match the distribution of the degree of misreporting, that is, the gap between the preference ranking of the observed second choice and the ex-ante/ex-post fair school derived from the expected/realized score distribution. If a student's simulated second choice is the same as the simulated first choice, then we re-generate the second choice randomly according to the same distribution until they are different. The 3rd-7th choices for each student are randomly assigned.

Scenario 4

Scenario 4 simulates the non-truth-telling students' second choice by considering its relation with the first choice after simulating the first choice as in Scenario 2. In particular, the second choices are assigned to match the distribution of the gap between the preference ranking of the first and the second choice observed in the data. The 3rd-7th choices for each student are randomly assigned.

Appendix B: Tables and Figures

Table 3: Payoff Table (with Student Scores and Rankings)

Student _id	Estimated /Realized Score	Score Rank	Payoff						
			College _A	College _B	College _C	College _D	College _E	College _F	College _G
1	121	20	65	80	25	35	55	10	45
2	116	25	80	65	45	25	10	35	55
3	111	30	65	80	10	45	25	55	35
4	137	4	80	65	35	45	25	10	55
5	132	9	65	80	55	45	25	10	35
6	109	32	80	65	10	45	25	55	35
7	136	5	65	80	45	10	25	55	35
8	127	14	80	65	55	35	45	10	25
9	123	18	65	80	25	45	55	35	10
10	119	22	80	65	55	10	35	45	25
11	124	17	65	80	10	55	35	45	25
12	105	36	80	65	10	55	25	45	35
13	128	13	65	80	45	55	25	10	35
14	135	6	80	65	55	45	25	35	10
15	108	33	65	80	45	35	25	10	55
16	130	11	80	65	45	10	55	25	35
17	131	10	65	80	45	35	55	10	25
18	106	35	80	65	45	25	10	35	55
19	138	3	65	80	10	35	45	55	25
20	113	28	80	65	45	10	25	35	55
21	140	1	65	80	45	35	25	10	55
22	114	27	80	65	55	25	35	45	10
23	110	31	65	80	45	35	10	55	25
24	112	29	80	65	25	35	55	10	45
25	126	15	65	80	45	35	25	10	55
26	120	21	80	65	45	55	25	35	10
27	129	12	65	80	45	35	55	10	25
28	133	8	80	65	25	45	35	10	55
29	139	2	65	80	35	55	25	10	45
30	115	26	80	65	25	10	35	55	45
31	122	19	65	80	10	55	35	45	25
32	134	7	80	65	10	45	55	25	35
33	125	16	65	80	55	25	35	10	45
34	117	24	80	65	35	45	10	55	25
35	107	34	65	80	55	10	25	35	45
36	118	23	80	65	10	55	45	35	25

Table 4: Determinants of Truth-telling: Logit Model

Independent Variable	Dependent Variable: Truth-telling		
	(1)	(2)	(3)
Boston	-0.5833***	-0.5823***	-0.5949***
(BOS_C-SD_C)	(0.0646)	(0.0646)	(0.0626)
Incomplete	-0.0463	-0.03539	-0.0475
(SD_I-SD_C)	(0.0696)	(0.0730)	(0.0720)
Boston*Incomplete	-0.05556	-0.0629	-0.0471
(BOS_I-BOS_C)-(SD_I-SD_C)	(0.0792)	(0.0805)	(0.0793)
Rank	-0.00344	-0.00336	-0.00263
	(0.00237)	(0.00239)	(0.00240)
Fair School Slots	-0.0181	-0.0191	-0.0241
	(0.0156)	(0.0158)	(0.0162)
Female		-0.000912	0.00214
		(0.0405)	(0.0405)
Age		-0.00154	0.000496
		(0.0147)	(0.0145)
Econ		0.0153	0.0245
		(0.0550)	(0.0548)
Engineer		0.0350	0.0506
		(0.0572)	(0.0574)
Science		-0.00267	-0.00205
		(0.0734)	(0.0735)
σ			0.208***
			(0.0763)
α			-0.0926
			(0.0810)
λ			0.00997
			(0.0124)
Observations	360	360	360
Pseudo R-squared	0.4022	0.4034	0.4186
BOS_I-BOS_C	-0.1019***	-0.0982***	-0.0946***
	(0.0377)	(0.0374)	(0.0362)
SD_I-BOS_C	0.5370***	0.5469***	0.5475***
	(0.0579)	(0.0594)	(0.0583)

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Coefficients report average marginal effects. λ takes the average value of its lower and upper bound. We also run regressions with the lower bound and upper bound of λ, and the results are similar.

Table 5: Determinants of Sorting: Logit Model

Independent Variable	Dependent Variable: Sorting					
	(1)	(2)	(3)	(4)	（5）	（6）
Boston	0.417***	0.421***	0.430***	0.417***	0.421***	0.429***
(BOS_C-SD_C)	(0.0751)	(0.0748)	(0.0741)	(0.0751)	(0.0749)	(0.0742)
Incomplete	-0.0602	-0.0557	-0.0496	-0.0602	-0.0562	-0.0507
(SD_I-SD_C)	(0.0620)	(0.0622)	(0.0623)	(0.0620)	(0.0623)	(0.0624)
Boston*Incomplete	0.0370	0.0266	0.0116	0.0370	0.0267	0.0127
(BOS_I-BOS_C)-(SD_I-SD_C)	(0.0956)	(0.0957)	(0.0957)	(0.0956)	(0.0957)	(0.0958)
Rank	-0.00225	-0.00209	-0.00205			
	(0.00299)	(0.00302)	(0.00302)			
Top 6 dummy				-0.0717	-0.0751	-0.0688
				(0.0937)	(0.0938)	(0.0937)
Fair School Slots	0.0133	0.0134	0.0111	-0.00880	-0.00865	-0.00967
	(0.0194)	(0.0194)	(0.0197)	(0.0218)	(0.0218)	(0.0219)
Female		0.00702	0.0105		0.00397	0.00758
		(0.0510)	(0.0515)		(0.0508)	(0.0512)
Age		-0.0122	-0.0126		-0.0116	-0.0118
		(0.0171)	(0.0171)		(0.0171)	(0.0171)
Econ		-0.0213	-0.00773		-0.0197	-0.00674
		(0.0925)	(0.0941)		(0.0922)	(0.0937)
Engineer		-0.0391	-0.0402		-0.0393	-0.0411
		(0.0951)	(0.0963)		(0.0946)	(0.0957)
Science		-0.106	-0.108		-0.113	-0.116
		(0.113)	(0.115)		(0.113)	(0.114)
σ			-0.111			-0.108
			(0.101)			(0.101)
α			0.154			0.153
			(0.104)			(0.105)
λ			-0.00967			-0.00951
			(0.0155)			(0.0155)
Observations	360	360	360	360	360	360
Pseudo R-squared	0.155	0.159	0.164	0.155	0.160	0.165
BOS_I-BOS_C	-0.0231	-0.0291	-0.0380	-0.0231	-0.0295	-0.0381
	(0.0728)	(0.0731)	(0.0728)	(0.0728)	(0.0731)	(0.0728)
SD_I-BOS_C	0.477***	0.477***	0.480***	0.477***	0.477***	0.480***
	(0.0670)	(0.0674)	(0.0670)	(0.0669)	(0.0673)	(0.0669)

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Coefficients report average marginal effects. λ takes the average value of its lower and upper bound. We also run regressions with the lower bound and upper bound of λ, and the results are similar.

Table 6: Fairness Consequences of Different Strategies

Panel A: Summary statistics of degree of mismatch

Mechanism	Report Strategy	# of Obs.	Mean	SD	Min	Max	T test for Mean=0 (p-value)
BOS_I	Truth-telling	1	-0.86	n.a.	-0.86	-0.86	n.a
	Sorting	68	-0.00044	0.5	-0.93	2	(0.9942)
	Other non-truth-telling	39	-0.69	1	-3	2	(0.0001)
BOS_C	Truth-telling	8	0.13	0.83	-1	2	(0.6845)
	Sorting	47	0.085	0.46	-1	2	(0.2093)
	Other non-truth-telling	17	-1.1	1	-3	1	(0.0006)
SD_I	Truth-telling	70	-0.011	0.58	-0.95	1.6	(0.8709)
	Sorting	19	-0.024	0.56	-0.84	1.9	(0.8559)
	Other non-truth-telling	19	-0.61	0.83	-2	0.74	(0.0046)
SD_C	Truth-telling	50	0.02	0.14	0	1	(0.3222)
	Sorting	17	0.12	0.49	0	2	(0.3322)
	Other non-truth-telling	5	-1	1	-2	0	(0.0890)

Panel B: Wilcoxon Rank-Sum Test of differences in degree of mismatch

Mechanism	Result
BOS_I	Sorting >Other≈ Truth-telling (p=0.0001) (p=0.8960)
BOS_C	Truth-telling ≈ Sorting> Other (p=0.7743) (p=0.0000)
SD_I	Truth-telling ≈ Sorting> Other (p=0.9680) (p=0.0411)
SD_C	Sorting >Other>Truth-telling (p=0.0020) (p=0.0000)

Panel C: F test of variance of degree of mismatch

Mechanism	Result
BOS_I	Other> Sorting (p=0.0000)
BOS_C	Other ≈ Truth-telling > Sorting (p=0.5923) (p=0.0061)
SD_I	Other >Truth-telling ≈ Sorting (p=0.0157) (p=0. 9485)
SD_C	Other ≈ Sorting >Truth-telling (p=0.156) (p=0.0000)

Note: Degree of mismatch for BOS_I and SD_I is the average value after 200 simulations generated from the score distribution.

Table 7: Efficiency Consequences of Different Strategies

Mechanism	Report Strategy	# of Obs.	Mean	SD	Min	Max
	Panel A: Summary statistics of relative payoff					
BOS_I	Truth-telling	1	0.866	.	0.866	0.866
	Sorting	68	1.011	0.119	0.855	1.571
	Other non-truth-telling	39	0.882	0.199	0.455	1.455
BOS_C	Truth-telling	8	1.038	0.177	0.846	1.455
	Sorting	47	1.025	0.124	0.818	1.571
	Other non-truth-telling	17	0.809	0.182	0.556	1.182
SD_I	Truth-telling	70	1.013	0.124	0.826	1.451
	Sorting	19	1.01	0.141	0.869	1.554
	Other non-truth-telling	19	0.89	0.156	0.628	1.168
SD_C	Truth-telling	50	1.004	0.031	1	1.222
	Sorting	17	1.034	0.139	1	1.571
	Other non-truth-telling	5	0.818	0.182	0.636	1

	Panel B: Wilcoxon Rank-Sum Test of differences in relative payoff
BOS_I	Sorting >Other ≈Truth-telling (p=0.0001) (p=0.8960)
BOS_C	Truth-telling≈ Sorting>Other (p=0.7575) (p=0.0000)
SD_I	Truth-telling ≈ Sorting>Other (p=0.7903) (p=0.0383)
SD_C	Sorting ≈Truth-telling>Other (p=0.4061) (p=0.0000)

	Panel C: F test of variance of relative payoff
BOS_I	Other >Sorting (p=0.0001)
BOS_C	Other ≈ Truth-telling > Sorting (p=0.9990) (p=0.0701)
SD_I	Other ≈ Sorting ≈ Truth-telling (p=0.6743) (p=0.4444)
SD_C	Other ≈ Sorting >Truth-telling (p=0.3891) (p=0.0000)

Note: The relative payoff is defined as a student's realized payoff divided by his payoff under the fair match. For BOS_I and SD_I, it is the average value after 200 simulations generated from the score distribution.

Table 8: Social Welfare Comparison Among Four Mechanisms

Welfare Measures	Panel A: Mean of Welfare Measures			
	BOS_I	BOS_C	SD_I	SD_C
Efficiency (by payoff per capita)	10.046	10.551	10.664	11.105
Ex-post fairness (by degree of mismatch)	1.032	0.642	0.551	0.266
Ex-post fairness (by number of blocking pairs)	0.948	0.651	0.349	0.145
Ex-ante fairness (by degree of mismatch)	0.963	0.876	0.749	0.598
Ex-ante fairness (by number of blocking pairs)	0.973	0.849	0.589	0.416

	Panel B: Wilcoxon Rank-Sum Test Result
Efficiency (by payoff per capita)	BOS_I < BOS_C < SD_I < SD_C (p=0.0000)(p=0.0000)(p=0.0000)
Ex-post fairness (by degree of mismatch)	BOS_I > BOS_C > SD_I > SD_C (p=0.0000)(p=0.0000)(p=0.0000)
Ex-post fairness (by number of blocking pairs)	BOS_I > BOS_C > SD_I > SD_C (p=0.0000)(p=0.0000)(p=0.0000)
Ex-ante fairness (by degree of mismatch)	BOS_I > BOS_C > SD_I > SD_C (p=0.0000)(p=0.0000)(p=0.0000)
Ex-ante fairness (by number of blocking pairs)	BOS_I > BOS_C > SD_I > SD_C (p=0.0000)(p=0.0000)(p=0.0000)

Table 9: The Effects of Changing Sorting Behavior under Boston Mechanisms

	Origin (1)	6% (2)	-6% (3)	12% (4)	-12% (5)	100% sorting (6)	100% truth-telling (7)
BOS_I							
Efficiency	10.046	10.249	9.897	10.429	9.74	11.431	11.611
Ex-post (un)fairness-1	1.032	0.949	1.09	0.869	1.165	0.435	0.157
Ex-post (un)fairness-2	0.948	0.894	1.001	0.84	1.066	0.638	0.071
Ex-ante (un)fairness-1	0.963	0.873	1.036	0.791	1.114	0.293	0.424
Ex-ante (un)fairness-2	0.973	0.898	1.022	0.825	1.103	0.541	0.262
Efficiency	All truth > all sorting > 12% > 6% > Origin > -6% > -12% (0.000) (0.000) (0.000) (0.000) (0.000) (0.000)						
Ex-post (un)fairness-1	-12% > -6% > Origin > 6% > 12% > all sorting > all truth (0.000) (0.000) (0.000) (0.000) (0.000) (0.000)						
Ex-post (un)fairness-2	-12% > -6% > Origin > 6% > 12% > all sorting > all truth (0.0001) (0.0008) (0.0004) (0.0022) (0.000) (0.000)						
Ex-ante (un)fairness-1	-12% > -6% > Origin > 6% > 12% > all truth > all sorting (0.000) (0.000) (0.000) (0.000) (0.000) (0.000)						
Ex-ante (un)fairness-2	-12% > -6% > Origin > 6% > 12% > all sorting > all truth (0.000) (0.0033) (0.0001) (0.0001) (0.000) (0.000)						
BOS_C							
Efficiency	10.551	10.713	10.336	10.893	10.11	11.509	11.61
Ex-post (un)fairness-1	0.642	0.557	0.744	0.478	0.849	0.169	0.16
Ex-post (un)fairness-2	0.651	0.587	0.753	0.519	0.86	0.42	0.073
Ex-ante (un)fairness-1	0.876	0.797	0.969	0.722	1.065	0.468	0.417
Ex-ante (un)fairness-2	0.849	0.782	0.944	0.709	1.047	0.568	0.253
Efficiency	All truth > all sorting > 12% > 6% > Origin > -6% > -12% (0.000) (0.000) (0.000) (0.000) (0.000) (0.000)						
Ex-post (un)fairness-1	-12% > -6% > Origin > 6% > 12% > all sorting ≈ all truth (0.000) (0.000) (0.000) (0.000) (0.000) (0.3045)						
Ex-post (un)fairness-2	-12% > -6% > Origin > 6% > 12% > all sorting > all truth (0.000) (0.000) (0.0015) (0.001) (0.000) (0.000)						
Ex-ante (un)fairness-1	-12% > -6% > Origin > 6% > 12% > all sorting > all truth (0.000) (0.000) (0.000) (0.000) (0.000) (0.000)						
Ex-ante (un)fairness-2	-12% > -6% > Origin > 6% > 12% > all sorting > all truth (0.000) (0.000) (0.0007) (0.0008) (0.000) (0.000)						

Note: Efficiency is measure by payoff per capita. Ex-ante or ex-post (un)fairness is measured by degree of mismatch (1) or the number of blocking pairs (2). Wilcoxon Rank-Sum Test is used to rank matching outcomes.

Table 10: The Effects of Changing Sorting Behavior under SD Mechanisms

	Origin （1）	6% （2）	-6% （3）	12% （4）	-12% （5）	100% sorting （6）	100% Truth-telling （7）
SD_I							
Efficiency	10.664	10.636	10.691	10.643	10.657	11.171	11.528
Ex-post (un)fairness-1	0.551	0.583	0.541	0.585	0.53	0.511	0
Ex-post (un)fairness-2	0.349	0.364	0.336	0.374	0.333	0.742	0
Ex-ante (un)fairness-1	0.749	0.742	0.735	0.733	0.765	0.409	0.407
Ex-ante (un)fairness-2	0.589	0.586	0.576	0.597	0.594	0.667	0.258
Efficiency	All truth > all sorting > -6% ≈ Origin ≈ -12% ≈ +12% ≈ +6% (0.000) (0.000) (0.323) (0.611) (0.604) (0.867)						
Ex-post (un)fairness-1	12% ≈ 6% > Origin ≈ -6% ≈ -12% ≈ all sorting > all truth (0.832) (0.017) (0.254) (0.522) (0.1225) (0.000)						
Ex-post (un)fairness-2	all sorting > 12% ≈ 6% > Origin ≈ -6% ≈ -12% > all truth (0.000) (0.339) (0.090) (0.104) (0.939)(0.000)						
Ex-ante (un)fairness-1	-12% ≈ Origin ≈ 6% ≈ -6% ≈ 12% > all sorting ≈ all truth (0.167) (0.659) (0.456) (0.755) (0.000)(0.306)						
Ex-ante (un)fairness-2	all sorting > 12% ≈ -12% ≈ Origin ≈ 6% ≈ -6% > all truth (0.000) (0.617) (0.826) (0.586) (0.490)(0.000)						
SD_C							
Efficiency	11.105	11.115	11.096	11.091	11.153	11.318	11.532
Ex-post (un)fairness-1	0.266	0.261	0.269	0.27	0.239	0.231	0
Ex-post (un)fairness-2	0.145	0.144	0.148	0.152	0.129	0.51	0
Ex-ante (un)fairness-1	0.598	0.581	0.596	0.603	0.577	0.557	0.398
Ex-ante (un)fairness-2	0.416	0.411	0.429	0.428	0.392	0.662	0.254
Efficiency	All truth > all sorting > -12% ≈ 6% ≈ Origin ≈ -6% ≈ 12% (0.000) (0.000) (0.206) (0.503) (0.688) (0.928)						
Ex-post (un)fairness-1	12% ≈ -6% ≈ Origin ≈ 6% ≈ -12% ≈ all sorting > all truth (0.887) (0.648) (0.718) (0.111) (0.294) (0.000)						
Ex-post (un)fairness-2	all sorting > 12% ≈ -6% ≈ Origin ≈ 6% > -12% > all truth (0.000) (0.606) (0.508) (0.911) (0.078)(0.000)						
Ex-ante (un)fairness-1	12% ≈ Origin ≈ -6% ≈ 6% ≈ -12% ≈ all sorting > all truth (0.899) (0.733) (0.164) (0.879) (0.121) (0.000)						
Ex-ante (un)fairness-2	all sorting > -6% ≈ 12% ≈ Origin ≈ 6% ≈ -12% > all truth (0.000) (0.859) (0.286) (0.634) (0.151)(0.000)						

Note: Efficiency is measured by payoff per capita. Ex-ante or ex-post (un)fairness is measured by degree of mismatch (1) or the number of blocking pairs (2). Wilcoxon Rank-Sum Test is used to rank matching outcomes.

Table 11: Determinants of Fair-Reporting: Logit Model

Independent Variable	Dependent Variable: Fair-Reporting		
	(1)	(2)	(3)
Boston	0.431***	0.439***	0.445***
(BOS_C-SD_C)	(0.0735)	(0.0730)	(0.0727)
Incomplete	-0.0278	-0.0280	-0.0258
(SD_I-SD_C)	(0.0587)	(0.0587)	(0.0589)
Boston*Incomplete	0.00463	-0.00726	-0.0174
(BOS_I-BOS_C)-(SD_I-SD_C)	(0.0942)	(0.0939)	(0.0941)
Rank	-0.00311	-0.00288	-0.00282
	(0.00296)	(0.00299)	(0.00300)
Fair School Slots	-0.000301	-0.00129	-0.00381
	(0.0192)	(0.0193)	(0.0195)
Female		0.0145	0.0190
		(0.0506)	(0.0510)
Age		0.00137	0.00154
		(0.0169)	(0.0169)
Econ		-0.0731	-0.0626
		(0.0903)	(0.0916)
Engineer		-0.108	-0.112
		(0.0928)	(0.0939)
Science		-0.0992	-0.106
		(0.111)	(0.112)
σ			-0.0631
			(0.0999)
α			0.134
			(0.104)
λ			-0.00981
			(0.0155)
Observations	360	360	360
Pseudo R-squared	0.158	0.162	0.166
BOS_I-BOS_C	-0.0231	-0.0353	-0.0432
	(0.0737)	(0.0737)	(0.0736)
SD_I-BOS_C	0.458***	0.467***	0.471***
	(0.0672)	(0.0669)	(0.0667)

Note: *** p<0.01, ** p<0.05, * p<0.1. Coefficients report average marginal effects. λ takes the average value of its lower and upper bound. We also run regressions with the lower bound and upper bound of λ, and the results are similar.

Table 12: Determinants of Misreport within Non-Truth-Telling

Independent Variable	Dependent Variable: Degree of Misreport					
	OLS			Ordered Probit		
	(1)	(2)	(3)	(4)	(5)	(6)
Boston	-0.0720	-0.0706	-0.153	-0.131	-0.130	-0.220
(BOS_C-SD_C)	(0.250)	(0.248)	(0.250)	(0.274)	(0.274)	(0.278)
Incomplete	0.116	0.128	0.0985	0.0354	0.0553	0.0114
(SD_I-SD_C)	(0.271)	(0.269)	(0.279)	(0.293)	(0.294)	(0.307)
Boston*Incomplete	0.0126	0.0115	0.101	0.0941	0.0864	0.176
(*BOS_I-BOS_C)-(SD_I-SD_C)*	(0.315)	(0.312)	(0.317)	(0.341)	(0.342)	(0.349)
Rank	-0.0171**	-0.0272***	-0.0264***	-0.0202***	-0.0331***	-0.0329***
	(0.00678)	(0.00838)	(0.00855)	(0.00754)	(0.00943)	(0.00968)
Fair School Slots		0.112**	0.104*		0.140**	0.134**
		(0.0550)	(0.0556)		(0.0608)	(0.0620)
Female			-0.0949			-0.0808
			(0.144)			(0.158)
Age			0.00718			0.00723
			(0.0480)			(0.0530)
Econ			0.425**			0.427**
			(0.189)			(0.210)
Engineer			0.326			0.330
			(0.198)			(0.219)
Science			0.0645			-0.0626
			(0.291)			(0.318)
σ			0.0227			0.103
			(0.300)			(0.330)
α			0.0209			-0.0181
			(0.321)			(0.353)
λ			-0.0663			-0.0737
			(0.0468)			(0.0514)
Observations	231	231	231	231	231	231
(Pseudo)R-squared	0.032	0.050	0.091	0.0141	0.0233	0.0394
BOS_I-BOS_C	0.1282	0.1394	0.1997	0.1295	0.1417	0.1870
	(0.1600)	(0.1590)	(0.1670)	(0.1743)	(0.1748)	(0.1852)
BOS_C-SD_I	-0.1875	-0.1985	-0.2513	-0.1665	-0.1854	-0.2317
	(0.2070)	(0.2056)	(0.2178)	(0.2231)	(0.2238)	(0.2395)

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. λ is the average value of its lower and upper bound. We also run regression with the lower bound and upper bound of λ, and the results are similar.

Table 13: OLS Regression of the Degree of Mismatch on Misreport

Independent Variable	Dependent Variable							
	Degree of Mismatch				Variance of Degree of Mismatch			
	BOS_I	BOS_C	SD_I	SD_C	BOS_I	BOS_C	SD_I	SD_C
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Truth-telling	-0.416	0.146	0.246*	0.0133	0.423	0.585**	0.211	0.0127
	(0.671)	(0.265)	(0.139)	(0.0841)	(0.703)	(0.240)	(0.139)	(0.0557)
Up-report	0.278*	0.341	0.397**	0.390**	0.894***	0.806***	0.421**	0.379***
	(0.152)	(0.242)	(0.196)	(0.157)	(0.159)	(0.219)	(0.196)	(0.104)
Down-report	-1.351***	-1.446***	-1.105***	-1.798***	0.118	0.0812	-0.0652	0.112
	(0.179)	(0.206)	(0.217)	(0.194)	(0.188)	(0.187)	(0.217)	(0.129)
Rank	0.00111	0.0228**	-0.0191***	0.00915**	0.0230***	0.00731	0.0260***	0.00651**
	(0.00824)	(0.00944)	(0.00680)	(0.00424)	(0.00864)	(0.00856)	(0.00681)	(0.00281)
Fair School	0.0899*	-0.0964	0.194***	-0.0343	-0.109**	0.0683	-0.109**	-0.0228
Slots	(0.0515)	(0.0634)	(0.0420)	(0.0288)	(0.0540)	(0.0574)	(0.0420)	(0.0191)
Female	-0.300**	0.283*	-0.143	0.0826	0.136	-0.304**	-0.0881	0.0792
	(0.130)	(0.160)	(0.101)	(0.0795)	(0.136)	(0.145)	(0.101)	(0.0527)
Age	0.0515	0.0516	0.0118	-0.0439*	-0.0991**	0.00231	0.00644	-0.0395**
	(0.0463)	(0.0457)	(0.0417)	(0.0242)	(0.0485)	(0.0414)	(0.0417)	(0.0160)
Econ	0.225	0.195	0.0157	0.0600	-0.176	-0.165	-0.0611	0.0228
	(0.156)		(0.135)	(0.114)	(0.164)	(0.289)	(0.135)	(0.0756)
Engineer	0.00339	-0.00320	0.0549	0.194	0.282	-0.185	-0.0122	0.141
	(0.163)	(0.308)	(0.141)	(0.128)	(0.171)	(0.279)	(0.141)	(0.0850)
Science	0.660*	0.100	0.207	0.112	0.0133	-0.235	-0.207	0.0734
	(0.376)	(0.350)	(0.200)	(0.150)	(0.394)	(0.317)	(0.200)	(0.0991)
σ	-0.423	0.221	-0.328	-0.00749	0.0407	0.333	-0.147	0.00232
	(0.269)	(0.367)	(0.200)	(0.141)	(0.282)	(0.333)	(0.201)	(0.0936)
α	0.667**	0.737**	0.0102	-0.140	-0.121	-0.459	-0.0838	-0.0915
	(0.292)	(0.354)	(0.218)	(0.145)	(0.305)	(0.321)	(0.218)	(0.0963)
λ	0.0323	-0.0284	-0.0580	0.00658	-0.0287	0.0565	0.00229	0.00529
	(0.0389)	(0.0548)	(0.0361)	(0.0216)	(0.0408)	(0.0497)	(0.0361)	(0.0143)
Constant	-1.935*	-1.775*	-0.826	0.861	2.357**	-0.218	0.295	0.762**
	(1.071)	(0.985)	(0.928)	(0.523)	(1.122)	(0.892)	(0.929)	(0.347)
Observations	108	72	108	72	108	72	108	72
R-squared	0.498	0.641	0.516	0.713	0.371	0.339	0.199	0.418

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in parentheses. λ is the average value of its lower and upper bound. We also run regression with the lower bound and upper bound of λ, and the results are similar. Degree of mismatch for BOS_I and SD_I is the average value after 200 simulations generated from the score distribution.

Fig. 8: The Effect of Changing the Proportion of Sorting Behavior under Boston Mechanisms

(Sorting behavior is measured by misreport)

Fig. 9: The Effect of Changing the Proportion of Sorting Behavior under SD Mechanisms

(sorting behavior is measured by misreport)

Appendix C: Instruction Manuals

The following are four instruction manuals for our experiments of four environments: Boston mechanism under incomplete information (BOS_I, or B1, as titled below), Boston mechanism under complete information (BOS_C, or B2), SD mechanism under incomplete information (SD_I, or S1) and SD mechanism under complete information (SD_C, or S2).

Instruction manual for each student differs in their estimated (for BOS/SD_I) or realized (for BOS/SD_C) scores, rankings and corresponding payoffs. Student payoffs are described in the "Student Payoff Table" in Table 3.

For each student, a survey is conducted to gather data on his/her personal information and risk attitude. The "Personal Background and Risk-Attitude Test Form" is also attached following the four instruction manuals.

1. Instructions - Mechanism B1

This is an experiment in the economics of decision making. The instructions are simple, and if you follow them carefully and make good decisions, you might earn a considerable amount of money. In this experiment, we simulate a procedure to allocate students to schools. The procedure, payment rules, and student allocation method are described below. Do not communicate with each other during the experiment. If you have questions at any point during the experiment, raise your hand and the experimenter will help you.

Procedure

- There are **36** students in this experiment. You are student**#1**.
- 36 school slots are available across seven schools. Each school slot is allocated to one student. There are three slots each at schools A and B, and six slots each at schools C, D, E, F and G. These schools differ in quality and specialties so different students will prefer difference schools. However, all students prefer schools A and B to all the other schools.
- Your payoff amount depends on the school slot you hold at the end of the experiment.
- You will receive 20 RMB yuan for participation, regardless of your school slot.

Payoff amounts in RMB yuan are outlined in the following table. These amounts reflect the desirability of the seven schools for you.

Slot received at school	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Payoff to you	65	80	55	35	45	25	10

The table is explained as follows:
– You will be paid 65 RMB yuan if you hold a slot at school A at the end of the experiment.

– You will be paid 80 RMB yuan if you hold a slot at school B at the end of the experiment.
– You will be paid 55 RMB yuan if you hold a slot at school C at the end of the experiment.
– You will be paid 35 RMB yuan if you hold a slot at school D at the end of the experiment.
– You will be paid 45 RMB yuan if you hold a slot at school E at the end of the experiment.
– You will be paid 25 RMB yuan if you hold a slot at school F at the end of the experiment.
– You will be paid 10 RMB yuan if you hold a slot at school G at the end of the experiment.

NOTE different students might have different payoff tables. That is, payoff by school might be different for different students.

- All students will be assigned a score. The details of the score assignment are discussed below.
- All schools have the same priority: they want the students with the highest scores.
- During the experiment, you first complete the Decision Sheet by indicating school preferences. Note that you need to rank all seven schools in order to indicate your preferences.
- After all participants have completed their Decision Sheets, we collect the sheets and start the allocation process.
- Once the allocations are determined, we inform each participant of his/her allocation slot and respective payoff.

Score assignment
One number is randomly picked from all integers from 105 to 140 and assigned as the **estimated score** for a randomly chosen student. Another number is picked from the remaining 35 integers for another student. This process continues until all 36 students have been assigned **estimated scores**. Note that no two students have the same **estimated score**. We will also inform each student with his/her rank among all the students.

Each student's **actual score** will not be assigned until his/her has listed his/her school choices on the Decision Sheet. Thus you will only know your **estimated score** when you make your school choices. The **actual score** can be up to **10** points higher or lower than the **estimated score**. For example, if the **estimated score** is 1**40**, then the **actual score** has equal probability of being any number from 1**30** to 1**50**. If an **actual score** assigned to a student has already been assigned to another student, then another score is randomly picked out of the possible range, until a different score is assigned to the student to be his/her **actual score**. Thus no two students have the same **actual score** also.

Allocation process

Once the **actual scores** are assigned, slots are allocated in seven rounds.

Round 1.
a. An application to the first ranked school in the Decision Sheet is sent for each student.
b. Each school accepts the students with the highest scores until all slots are filled. These students and their assignments are removed from the system. The remaining applications for each respective school are rejected.

Round 2.
a. The rejected applications are sent to his/her second ranked school in the Decision Sheet.
b. If a school still has available slots remaining from Round 1, then it accepts the students with the highest scores until all slots are filled. The remaining applications are rejected.

.
.
. ...

Round 6.
a. The application of each student who is rejected by his/her top five choices is sent to his/her sixth choice.
b. If a school still has slots available, then it accepts the students with higher scores until all slots are filled. The remaining applications are rejected.

Round 7.
Each remaining student is assigned a slot at his/her last choice.

An Example
We will go through a simple example to illustrate how the allocation method works.
Students and Schools. In the example, there are four students, 1-4, and 3 schools, A-C. There is one slot each at school A and B, 2 slots at school C.

Student ID Number: 1, 2, 3, 4	Schools: A, B, C

Score Assignment. Suppose the estimated score assigned to each student is as the following (rankings of those estimated scores are also shown):

	Estimated Score	Rank
Student 1	135	1
Student 2	123	2
Student 3	114	3
Student 4	107	4

Submitting School Rankings. The students submit the following school rankings:

	1st Choice	2nd Choice	3rd Choice
Student 1	A	B	C
Student 2	A	B	C
Student 3	B	A	C

Student 4	B	A	C

Allocation Process.
Suppose the actual score assigned to each student is realized as the following (ranks of those actual scores are also shown):

	Actual Score	Rank
Student 1	135	1
Student 2	128	2
Student 3	110	4
Student 4	112	3

Round 1. Each student applies to his/her first choice: Student 1 and 2 apply to school A; student 3 and 4 apply to school B.

- School A accepts Student 1 and rejects Student 2; School B accepts Student 4 and rejects Student 3.

Round 2. Each student rejected in Round 1 then applies to his/her second choice: Student 2 applies to School B, and Student 3 applies to School A.

- No slots are left at either School A or School B, so both Student 2 and Student 3 are rejected.

Round 3. Each student rejected in Round 2 then applies to his/her third choice: Both Student 2 and Student 3 apply to School C.

- School C (with 2 slots) accepts both Student 2 and Student 3.

Based on this method, the final allocations are:

Student	1	2	3	4
School	A	C	C	B

You will have 15 minutes to go over the instructions at your own pace, and make your decisions. Are there any questions?

Decision Sheet - Mechanism B1

• Recall: You are student #1.
• Recall: Your payoff amount depends on the school slot you hold at the end of the experiment.
……
• Your **estimated score** is **121**. This score ranks no **20** among all the 36 students.

Please write down your ranking of the schools (A through G) from your first choice to your last choice. Please rank ALL seven schools.

1st Choice	2nd Choice	3rd Choice	4th Choice	5th Choice	6th Choice	Last Choice

Your I.D:

Your Name (print):

This is the end of the experiment. Please go on to fill out the following short form concerning your personal background and risk attitudes. The form is important for analyzing the experimental results. After you finish the form, please remain seated until the experimenter collects your Decision Sheet. The experimenter will inform each student of his/her allocation slot and respective payoff once it is computed.

2. Instructions - Mechanism B2

……

Score assignment

One number is randomly picked from all the integers from 105 to 140 and assigned as the score for a randomly chosen student. Another number is picked from the remaining 35 integers for another student. This process continues until all 36 students have been assigned scores. Note that no two students have the same score. We will also inform each student with his/her rank among all the students.

Allocation process

Once the scores are assigned, and all the students submit their school preferences, slots are allocated in seven rounds.

Round 1.

……

Round 7.

……

An Example

We will go through a simple example to illustrate how the allocation method works.
Students and Schools. In the example, there are four students, 1-4, and 3 schools, A-C. There is one slot each at school A and B, 2 slots at school C.

Student ID Number: 1, 2, 3, 4	Schools: A, B, C

Score Assignment. Suppose the score assigned to each student is as the following (rankings of those scores are also shown):

	Actual Score	Rank
Student 1	135	1
Student 2	123	2
Student 3	114	3
Student 4	107	4

Submitting School Rankings. The students submit the following school rankings:

	1st Choice	2nd Choice	3rd Choice
Student 1	A	B	C

Student 2	A	B	C
Student 3	B	A	C
Student 4	B	A	C

Allocation Process.
Round 1. Each student applies to his/her first choice: Student 1 and 2 apply to school A; student 3 and 4 apply to school B.

- School A accepts Student 1 and rejects Student 2; School B accepts Student 3 and rejects Student 4.

Round 2. Each student rejected in Round 1 then applies to his/her second choice: Student 2 applies to School B, and Student 4 applies to School A.

- No slots are left at either School A or School B, so both Student 2 and Student 4 are rejected.

Round 3. Each student rejected in Round 2 then applies to his/her third choice: Both Student 2 and Student 4 apply to School C.

- School C (with 2 slots) accepts both Student 2 and Student 4.

Based on this method, the final allocations are:

Student	1	2	3	4
School	A	C	B	C

You will have 15 minutes to go over the instructions at your own pace, and make your decisions. Are there any questions?

Decision Sheet - Mechanism B2

……

3. Instructions - Mechanism S1

……

Score assignment

……

Allocation process

Once the **actual scores** are assigned, slots are allocated in thirty-six rounds.

Round 1.
The student with the highest score receives a slot at his/her top choice school.

Round 2.
The student with the second highest score is assigned her top choice among the remaining slots.

.

.

. ...

Round 36.

The student with the lowest score is assigned to the last slot.

An Example
We will go through a simple example to illustrate how the allocation method works.
Students and Schools. In the example, there are four students, 1-4, and 3 schools, A-C. There is one slot each at school A and B, 2 slots at school C.

Student ID Number: 1, 2, 3, 4	Schools: A, B, C

Score Assignment. Suppose the estimated score assigned to each student is as the following (rankings of those estimated scores are also shown):

	Estimated Score	Rank
Student 1	135	1
Student 2	123	2
Student 3	114	3
Student 4	107	4

Submitting School Rankings. The students submit the following school rankings:

	1st Choice	2nd Choice	3rd Choice
Student 1	A	B	C
Student 2	A	B	C
Student 3	B	A	C
Student 4	B	A	C

Allocation Process.
Suppose the actual score assigned to each student is realized as the following (ranks of those actual scores are also shown):

	Actual Score	Rank
Student 1	135	1
Student 2	128	2
Student 3	110	4
Student 4	112	3

Round 1. The student with the highest score receives a slot at his/her top choice school.

- Student 1 receives a slot at school A.

Round 2. The student with the second highest score is assigned her top choice among the remaining slots.

- Student 2 receives a slot at school B. (She cannot receive a slot at school A, her top choice school, because the slot has been taken by student 1.)

Round 3. The student with the third highest score is assigned her top choice among the remaining slots.

- Student 4 receives a slot at school C. (Slots at school A and B have been taken by student 1 and 2.)

Round 4. The student with the lowest score is assigned to the last slot.

- Student 3 receives a slot at school C.

Based on this method, the final allocations are:

Student	1	2	3	4
School	A	B	C	C

You will have 15 minutes to go over the instructions at your own pace, and make your decisions. Are there any questions?

Decision Sheet - Mechanism S1

……

4. Instructions - Mechanism S2

……

Score assignment

One number is randomly picked from all the integers from 105 to 140 and assigned as the score for a randomly chosen student. Another number is picked from the remaining 35 integers for another student. This process continues until all 36 students have been assigned scores. Note that no two students have the same score.

Allocation process

Once the scores are assigned, slots are allocated in thirty-six rounds.

Round 1.

. ...

Round 36.

……

An Example

We will go through a simple example to illustrate how the allocation method works. **Students and Schools.** In the example, there are four students, 1-4, and 3 schools, A-C. There is one slot each at school A and B, 2 slots at school C.

Student ID Number: 1, 2, 3, 4	Schools: A, B, C

Score Assignment. Suppose the score assigned to each student is as the following (rankings of those scores are also shown):

	Score	Rank
Student 1	135	1
Student 2	123	2
Student 3	114	3
Student 4	107	4

Submitting School Rankings. The students submit the following school rankings:

	1st Choice	2nd Choice	3rd Choice
Student 1	A	B	C
Student 2	A	B	C
Student 3	B	A	C
Student 4	B	A	C

Allocation Process.
Round 1. The student with the highest score receives a slot at his/her top choice school.

- Student 1 receives a slot at school A.

Round 2. The student with the second highest score is assigned her top choice among the remaining slots.

- Student 2 receives a slot at school B. (She cannot receive a slot at school A, her top choice school, because the slot has been taken by student 1.)

Round 3. The student with the third highest score is assigned her top choice among the remaining slots.

- Student 3 receives a slot at school C. (Slots at school A and B have been taken by student 1 and 2.)

Round 4. The student with the lowest score is assigned to the last slot.

- Student 4 receives a slot at school C.

Based on this method, the final allocations are:

Student	1	2	3	4
School	A	B	C	C

You will have 15 minutes to go over the instructions at your own pace, and make your decisions. Are there any questions?

Decision Sheet - Mechanism S2
……

5. Personal Background and Risk-Attitude Test Form

1. ……
2. Your gender is ___ (F/M).
3. Your age is ___
4. You major is ___
5. In what province and which year did you take you college entrance exam?___ (If you did not take the college entrance exam, please indicate the province your high school is in, and the year you graduate from high school.)
6. The following is a risk attitude test form. Please continue.

NOTE to pay for doing this risk attitude test form, we will randomly choose one participant in each experimental session. Then for the chosen participant we will randomly choose one row among all the 35 rows in three tables below. The chosen participant would be paid according to his/her lottery chosen at that row. The lottery would be drawn publicly on the spot. The unit of payment for this test is 0.5RMB. Good Luck!

You are going to choose from two lotteries, A and B, whose outcome will be determined by a random draw of 10 balls in a cage, with the balls being numbered 1, 2, 3,....10. For Table 1, at which row of lottery pairs would you begin to accept the Lottery B over Lottery A?

Table 1

Row	Lottery A		Lottery B	
	Ball 1-3	Ball 4-10	Ball 1	Ball 2-10
1	40	10	68	5
2	40	10	75	5
3	40	10	83	5
4	40	10	93	5
5	40	10	106	5
6	40	10	125	5
7	40	10	150	5
8	40	10	185	5
9	40	10	220	5
10	40	10	300	5
11	40	10	400	5
12	40	10	600	5
13	40	10	1,000	5
14	40	10	1,700	5

Your answer is in table 1:
I choose Lottery A for Row 1 to , and Lottery B for Row to 14.
Now consider another pair of Lotteries, C and D. Now for Table 2, at which row of lottery pairs would you begin to accept the Lottery D over Lottery C?

Table 2

Row	Lottery C		Lottery D	
	Ball 1-9	Ball 10	Ball 1-7	Ball 8-10
1	40	30	54	5
2	40	30	56	5
3	40	30	58	5
4	40	30	60	5
5	40	30	62	5
6	40	30	65	5
7	40	30	68	5
8	40	30	72	5
9	40	30	77	5
10	40	30	83	5
11	40	30	90	5
12	40	30	100	5
13	40	30	110	5

14	40	30	130	5

Your answer is in table 2:
I choose Lottery C for Row 1 to , and Lottery D for Row to 14.

Consider the final pair of Lotteries, E and F. Now for Table 3, at which row of lottery pairs would you begin to accept the Lottery F over Lottery E? (Note: Negative income implies money you are going to lose.)

Table 3

Row	Lottery E		Lottery F	
	Ball 1-5	Ball 6-10	Ball 1-5	Ball 6-10
1	25	-4	30	-21
2	4	-4	30	-21
3	1	-4	30	-21
4	1	-4	30	-16
5	1	-8	30	-16
6	1	-8	30	-14
7	1	-8	30	-11

Your answer is in table 3:
I choose Lottery E for Row 1 to , and Lottery F for Row to 14.