WINPEC Working Paper Series No.E1913
August 2019

Strategy-proofness in experimental matching markets

Pablo Guillen and Robert F. Veszteg

Waseda INstitute of Political EConomy
Waseda University
Tokyo, Japan

Strategy-proofness in experimental matching markets

Pablo Guillen[*]
The University of Sydney

Róbert F. Veszteg[†]
Waseda University

September 21, 2019

Abstract

We introduce two novel matching mechanisms, Reverse Top Trading Cycles (RTTC) and Reverse Deferred Acceptance (RDA), with the purpose of challenging the idea that the theoretical property of strategy-proofness induces high rates of truth-telling in economic experiments. RTTC and RDA are identical to the celebrated Top Trading Cycles (TTC) and Deferred Acceptance (DA) mechanisms, respectively, in all their theoretical properties except that their dominant-strategy equilibrium is to report one's preferences in the order opposite to the way they were induced. With the focal truth-telling strategy being out of equilibrium, we are able to perform a clear measurement of how much of the truth-telling reported for strategy-proof mechanisms is compatible with rational behavior and how much of it is caused by confused decision-makers following a default (very focal) strategy without understanding the structure of the game. In a school-allocation setting, we find that roughly half of the observed truth-telling under TTC and DA is the result of naïve (non-strategic) behavior. Only 13-29% of participants' actions in RTTC and RDA are compatible with rational behavior. Further than that, by looking at the responses of those seemingly rational participants in control tasks, it becomes clear that even them lack a basic understanding of the game incentives. We argue that the use of a default option, confusion and other behavioral biases account for the vast majority of truthful play in both TTC and DA in laboratory experiments.

Keywords: matching; strategy-proofness; truth-telling; focal point; rationality; laboratory experiment; school choice; revelation principle

JEL Classification Numbers: C78, D47, C91.

[*]Faculty of Economics, Room 340 Merewether Building (H04), NSW 2006 Sydney, Australia; pablo.guillen@sydney.edu.au

[†]School of Political Science and Economics, 1-6-1 Nishiwaseda, Shinjuku-ku, 169-8050 Tokyo, Japan; rveszteg@waseda.jp

1 Introduction

Laboratory experiments have been instrumental for the success of market design.[1] This is true, in particular, for centralized markets based on matching mechanisms. Those experiments show how strategy-proof (i.e., dominant-strategy incentive-compatible) mechanisms such as Deferred Acceptance (DA) and Top Trading Cycles (TTC) outperform non-strategy-proof mechanisms like the so-called Boston Mechanism (BOS). That is, both DA and TTC induce higher truth-telling rates and efficiency than BOS. This experimental evidence has been used as the "smoking gun" to convince stakeholders to adopt DA or TTC. The canonical reference for our claim is the famous Abdulkadiroğlu et al. (2006) paper which reports the efforts of a group of researchers to convince the Boston Public Schools planning team to adopt a strategy-proof matching mechanism. The researchers gathered theoretical and empirical, both historical and experimental, evidence about the vulnerability of the old (BOS) mechanism to preference misrepresentation (Abdulkadiroğlu and Sönmez, 2003; Chen and Sönmez, 2006; Roth 1991). Among those, Chen and Sönmez (2006) offers the seminal comparison of the strategy-proof DA and TTC with the non-strategy-proof BOS. In that experiment, both DA and TTC generate more truthtelling than BOS and also outperform it in terms of efficiency. Given the modest truth-telling rates observed in DA and TTC, between 43%-50% and 56%-72% respectively, Chen and Sönmez (2006) is cautious in terms of promoting the virtues of theoretically strategy-proof mechanisms and recommends participants to be instructed so that they would not hurt themselves by misrepresenting their preferences. The authors seem, however, to assume that truth-telling participants understand the incentives. Since Chen and Sönmez (2006), other articles have reported different and, generally higher, truth-telling rates (Calsamiglia et al., 2010: DA 57%-58%, TTC 62%-74%; Pais and Pintér, 2008: DA 67%-82%, TTC 87%-96%; Pais et al., 2011: DA 58%-76%, TTC 62%-84%). This evidence did, perhaps, prompted overenthusiastic comments such as the following one in Pathak and Sönmez (2013): "Another factor [in favour of using strategy-proof mechanisms] was the potential to use unmanipulated preference data generated by the student assignment mechanism in various policy-related issues including the evaluation of schools".[2]

More recently, a handful or articles have been trying to evaluate the perfor-

[1] Roth (2015) offers a concise summary of experimental research related to market design.

[2] For a much less selective and concise review of experimental findings on matching markets, refer to Hakimov and Kübler (2019).

mance of strategy-proof mechanisms in the field and in the laboratory. Rees-Jones and Skowronek (2018) reports the results of an online survey run with recent participants in the National Residency Matching Program (NRMP) in which 23% of participants fail to play the dominant strategy. Note that, as compared to the typical undergraduate college students sampled for a laboratory experiment, student participants in NRMP have high cognitive ability, they have access to excellent resources in terms of high quality advice from the NRMP itself, and they face a game with extremely high stakes. Similar results are reported by Rees-Jones (2018), Hassidim et al. (2017, 2018) and Shorrer and Sóvágó (2018). Notwithstanding the significant proportion of participants deviating from truthful preference revelation in those studies, they do not provide much indication on whether the remaining majority (which did send truthful messages) does indeed understand the incentives of the game. A more careful approach taken by Guillen and Hakimov (2017) uses a within-subject experimental design based on TTC in a tightly controlled environment in which decision-makers report their preferences both with and without being informed about others' choice. Their data reveal that most participants (69%) fail to adapt to changing circumstances and to play the dominant strategy. Only 31% of them exhibit behavior that is compatible with theory.

We run an experiment to investigate the robustness of the commonly reported high truth-telling rates of 60-80% in matching laboratory experiments. That is, our research goal is to determine whether the majority of participants tell the truth because they understand their incentives to do so or because they simply follow a default and choose a salient strategy.

Indeed, the essential problem with matching experiments lies in the complexity of games under study and that the induced preference order constitutes a strong focal point. These two are not independent, because the induced preference order can be understood as a default from which experimental subjects may or may not decide to depart. Simply put, an unsophisticated experimental subject who does not understand the strategic environment induced by the experimenter and the incentives of the experiment may just submit the induced preference order and unwittingly play the dominant strategy of the game. That could be interpreted as supporting evidence to the underlying theoretical model. In the same vein, a pseudo-sophisticated subject may be more inclined to manipulate the somewhat more transparent DA than the arguably obscure TTC, thus explaining the higher truth-telling rates observed for TTC when compared to DA.[3] Experimental stud-

[3]For the relative obscurity of TTC, refer to the comments in Pathak (2017) about the aversion of

ies by Guillen and Hing (2014) and Guillen and Hakimov (2018) show that the standard experimental instructions explaining matching algorithms are too difficult to understand and that the majority of experimental participants can be easily influenced by (correct or incorrect) advice.

To some extent, the literature shows awareness of the bias introduced by unsophisticated players to the desired empirical test of theoretical models and to the interpretation of the experimental results. Experimental studies targeting complex theories and mechanisms like centralized matching markets often rely on protocols that include both a solved numerical example and perhaps an incentive-based quiz (i.e, an example to be solved by participants) prior to the main part of the experiment. The authors of these articles typically claim that participants who manage to solve the example do understand the incentives and are able to figure out that the induced game has an equilibrium in dominant strategies. This, however, is not a convincing argument as such incentive-based quizzes only test a cursory understanding of the instructions rather than the understanding of the game incentives and thus strategy-proofness in DA or TTC.[4]

Additionally, many matching experiments set a top priority for the second-best object (school) thus trying to temp subjects to manipulate the mechanism by inducing a so-called district-school bias (DSB). They do so, because playing the dominant strategy under these circumstances may well be understood as a good understanding of the incentives. Once again, we are skeptical about this claim. In this paper, we report results from a carefully designed experiment (with the above-mentioned preference structure) to distinguish between subjects who play the dominant strategy because it is a default (and a strong focal point) from subjects who play it for other, perhaps rational strategic, reasons.

Our design includes two baseline treatments, one for DA and another for TTC, that follow the standards set by Chen and Sönmez (2006). Our experiments therefore study one-shot interaction, are based on an induced priority order, use standard instructions explaining the workings on the corresponding algorithm, and incorporate a solved example and an incentive-based example, or quiz, to be solved by participants. Also, every participant in our experiments faces DSB. Not surprisingly, truth-telling rates in our two baseline treatments are high and in line with well-known results from the existing literature.

Our treatments introduce two novel matching mechanisms, Reverse DA (RDA)

school boards to adopting it.

[4]See the online appendix for the standard quiz used in this and other studies.

and Reverse TTC (RTTC), as benchmarks. Both are small variations of DA and TTC, respectively. The idea is that the central clearinghouse runs an algorithm that sends out proposals in the reverse order, starting from the lowest ranked object moving gradually towards the highest ranked. It took merely a few word changes in the instructions of DA and TTC to describe and induce the reversed mechanisms. Namely, "highest" got replaced by "lowest".[5] This change results in RDA and RTTC to have a dominant-strategy equilibrium in which participants are to submit their preferences in exactly the opposite order they are induced. Our benchmark mechanisms are not strategy-proof, as participants have incentives to lie, but they are dominant-strategy incentive-compatible and otherwise identical in other desirable and celebrated theoretical properties to DA and TTC.

Note that when comparing DA to RDA and TTC to RTTC our design is more parsimonious than previous experimental tests of matching mechanisms which simply compare the performance of very different mechanisms. In particular, it is easier to claim for us that DA and RDA, on one hand, and TTC and RTTC on the other, are more comparable in terms of complexity than, for instance, DA and TTC or in the most extreme case BOS and TTC. The rate at which participants manage to solve the quiz across our four treatments corroborates this claim.[6]

In a nutshell, we find that the majority of participants fall in one or another behavioral trap. Only 16% and 26% of them play the dominant strategy in RDA and RTTC, down from 68% and 46% in DA and TTC. The induced preference order is played by 31% and 22% in RDA and RTTC, while DSB and what we call a naïve district school bias (nDSB) accounts for the majority of other observed strategies. We observe that truth-telling is not only the modal strategy for DA and TTC, but it remains so even under RDA and ranks as the second most-frequently played strategy under RTTC. Many decision-makers stick to the truth not because it constitutes an optimal strategy, but because it is the default strategy in a very complex situation. Our results indicate that, by far, default play is the main driver of the *usual* experimental results. The complexity of matching mechanism obscures strategy-proofness to the extent that it becomes an irrelevant theoretical property in laboratory experiments.

[5] For instance, "an application to the highest ranked school" was replaced by "an application to the lowest ranked school". Refer to the experimental instructions in section B of the appendix for details.

[6] Our design also reduces the bias that uncontrolled other-regarding preferences could create by observing human behavior in interaction with automated opponents. Refer to section 2 and to the experimental instructions in section B of the appendix for details.

On a positive note, our results also indicate that truth-telling, being such a salient strategy, may well be also salient in the field. That is, real-world participants who do not know how to behave strategically may just choose to tell the truth. However, this optimistic idea should be taken with some caution as it questions fundamental assumptions behind theoretical models and it undermines the very purpose of the theory behind market design.

The design of our experiment and the procedures used are thoroughly explained in section 2. Section 3 goes over the theoretical prediction and the experimental results, and section 4 discusses the implication of the results and concludes. The appendix, available online, describes the deferred-acceptance and the top-trading-cycles algorithms and includes the experimental instructions, examples and quizzes used.

2 Experimental design

Our data were collected through 10 computerized sessions (using zTree; Fischbacher, 2007) at Waseda University (Tokyo, Japan) between December 2017 and January 2018.[7] A total of 209 students participated in our sessions (DA: 50, RDA: 49, TTC: 56, RTTC: 54). Sessions lasted approximately for an hour. Participants were paid a fixed show-up fee of ¥700 and an additional ¥913 on average based on their performance.[8] No one participated in more than one session (between-subject design).[9]

Upon arrival to the experimental laboratory, participants were randomly assigned to a computer terminal and received all relevant instructions in written form (in Japanese). We kept our instructions as close as possible to the ones used by Chen and Sönmez (2006) and numerous follow-up studies.[10] The written text in-

[7] Waseda University is one of the top private universities in Japan. Admission is highly selective and depends both on high-school marks and results from an entry exam administered by the university. Also, note that Japanese high schools are of an extremely high quality. They ranked #5 in the world in the last PISA international mathematics comparisons, where South Korea ranked #6 and the USA ranked #40. In conclusion, the participants in our sample are much better trained in mathematics than the ones in most previous studies.

[8] Around the time of our experiments ¥1000 were equivalent to around $9, and would be enough to buy two lunch boxes on campus.

[9] All our participants are volunteers who signed up for the experiment by responding to an online advertisement. They are Waseda-University students of various majors. Our average participant is 21 years old, and 60% of our participants are male.

[10] In July 2019, Scopus listed 99 citations for our main references (Chen and Sönmez, 2006) whose experimental design follows the above-described pattern. 40% of the citing documents report new experimental results and almost 10% of them rely on a very similar and often nearly identical

formed participants about procedural matters, contained the detailed description of the matching algorithm used by the clearing house and also an illustrative example for how the algorithm works.

Each treatment (i.e., matching mechanism) was implemented in a different experimental session. During the experiment, after reading the instructions, each participant first completed a quiz which consisted of an allocation instance.[11] Subsequently, participants considered a single school-allocation task and submitted a preference ordering to the clearing house based on which the final matching and payoffs were determined.[12]

The (main) school-allocation task was a matching problem involving four schools with one vacant seat each and four students. Each participant played the role of a student (H), was assigned to a different matching problem, and interacted with three computers in the role of the other three students (R_1, R_2, R_3). Participants were told, in the instructions, that computers would act to maximize their expected gain. These design features were chosen to increase experimental control.[13]

We implemented the same school-allocation problem in all markets in all sessions. The columns in table 1 show students' preferences by listing schools from best to worst. Participants in the experiment could earn ¥1000 when matched to their favorite school, ¥500 when matched to the second-best, ¥200 when matched to the third-best, and ¥50 when matched to the least favorite school. Priorities for schools are represented by frames in the table: students enjoy priority over all others at schools whose letter appears in a frame. All remaining priorities were determined with a fair lottery by the matching mechanism.

The sections in our experimental instructions that described the matching algorithm and illustrated its rules through a numerical example are essentially identical to those used by Chen et al. (2016), thus comparable to many related experimental studies of the school-choice problem.[14] Descriptions of the reverse algorithms dif-

design. As for the instructions, we adapted the text as published by Chen et al. (2016) to our algorithms. All instructions, both in Japanese and English, are available from the authors upon request. A sample is to be found in section B of the appendix.

[11]Participants had 10 minutes to solve the quiz and were paid ¥100 for a correct solution. Given that the quiz aims at measuring participants' understanding of the analyzed strategic interaction, participants could only submit one solution and were not informed whether that was correct or not.

[12]Participants had up to 20 minutes to complete the school-allocation task. Just like in Chen and Sönmez (2006) and the relevant related literature, the allocation task was not repeated.

[13]The chosen design features not only make belief elicitation about the other decision-makers' rationality and behavior unnecessary, but they also strengthen the induced-value method in implementing the desired situation. Note that other-regarding considerations are unlikely to influence behavior when the *others* are simply computer codes.

[14]The only notable change in the instructions, apart from the language, is that names for schools

STUDENTS				MONETARY
R_1	R_2	R_3	H	PAYOFF
[C]	B	[D]	A	¥1000
A	[A]	C	[B]	¥500
B	C	A	C	¥200
D	D	B	D	¥50

Table 1: Preference profile for students (R_1, R_2, R_3, H) and priorities for schools (A, B, C, D). Students enjoy priority over all the others at schools marked by frame.

fer only in a few words from those of the celebrated DA and TTC algorithms. In particular, the word *highest* was replaced by *lowest* in the sentence "an application to the highest ranked school on the submitted ranking is sent for each participant", etc. Similarly, the overlap between the illustrating examples and quiz questions for the direct and the reverse algorithms was almost 100%. Recall that each participant in our experiment faced only one school-choice problem and therefore only one matching algorithm. Thus, the similarity among instructions used across treatments aimed at increased experimental control without causing unnecessary confusion to participants.

Participants were required to make a single decision and submit a preference list to the clearing house which determined the final matching based on the received lists and the announced algorithm. Participants had to make a decision in a partial-information setting, meaning that they only held precise information about their own preferences (column H in table 1). As for the rest, the written instructions stated that "[d]ifferent participants and different computers may or may not have different payoff tables. That is, payoff by school may or may not be different for different participants and different computers."

Finally, before they were paid individually and privately, participants were asked to fill out a questionnaire that included the typical questions on demographics and some others specific to the school-allocation task considered.

---

in the explanatory example were replaced by neutral symbols that are often used in generic lists in Japanese.

3 Results

The decision-problem implemented by our experimental design constitutes an instance of the so-called school-choice problem (Abdulkadiroğlu and Sönmez, 2003). Decision-makers belong to one of two disjoint sets (students or schools). Each student has strict preference over all schools. Each school has a maximum capacity and a strict priority ordering over all students. A solution to the school-choice problem is a matching, that is an assignment of schools to students.

3.1 Theoretical predictions

A matching is (Pareto) *efficient* if there is no other matching which assigns each student a weakly better school and at least one student a strictly better school. A matching is said to eliminate *justified envy* if there are no blocking pairs; in other words, there does not exist any student-school pair who are not matched to each other but would prefer to be so matched.[15] A *stable* matching eliminates justified envy, and it is individually rational; that is, each agent is matched to someone whom she finds better than being unmatched. Achieving a stable matching constitutes the goal of the theoretical literature (and the matchmaker's).[16]

The matching literature is primarily concerned with centralized solutions to the school-choice problem. In such a centralized matching market each decision-maker submits a list of preferences to the matchmaker (or a central clearinghouse) who produces a matching by processing all lists by means of a matching algorithm. Matching algorithms are simply the rules followed by the matchmaker to produce the final matching. Those rules are assumed to be known by all decision-makers involved. Note that in the school-choice problem only one side of the market, students, are assumed to be strategic. Schools' priorities are known by the matchmaker, or are reported truthfully to her.[17]

The above centralized solution is called *matching mechanism*. It is a situation of strategic interdependence in which the final outcome is jointly determined based

[15]The above definitions are widely accepted and used in the literature. Note that efficiency is a one-sided concept given that it only takes students' preferences into consideration. Justified envy, however, is two-sided: both students' and schools' preferences are considered.

[16]Besides the desirability of a stable matching from a normative point of view, stability as a prediction or solution to the school-allocation problem sounds reasonable when decision-makers are allowed to trade and switch partners in a frictionless decentralized way after the matchmaker has announced the outcome. Then, by definition, we should not expect unstable matchings to survive, only stable ones could prevail.

[17]The literature refers to the problem that involves strategizing school as the *college-admission* problem.

on the strategies chosen by the decision-makers (preference lists submitted by students), the matching algorithm and some other fixed parameters (school priorities).

In our experiment, students were first tentatively assigned to their *local schools* where they enjoyed priority over all the other students and then were required to submit their rankings over the four schools. Note that our example is such that each student enjoys priority at a different school. Students R_1 and R_3 ranked their local schools above all the others, while the remaining two students R_2 and H were in conflict with each other for the other two schools. This preference profile creates one of the simplest non-obvious school-allocation problems. The student-stable and also efficient matching in this problem is $\{(R_1, C), (R_2, B), (R_3, D), (H, A)\}$.[18]

The experimental treatments differ from each other in terms of the algorithm used by the matchmaker. The DA and TTC treatments operate with two well-known algorithms: the deferred-acceptance algorithm (Gale and Shapley, 1962) and the top-trading-cycles algorithm (Shapley and Scarf, 1974), respectively.[19] The other two algorithms, RDA and RTTC, are the reverse versions of the above two: their rules are identical to "original" algorithm except that proposals are first sent to the lowest (instead oh highest) ranked partner on the submitted ranking, then to the one above (instead of below), etc.

A matching mechanism is said to be *efficient* if it always selects a Pareto efficient matching, *stable* if it always selects a stable matching, *incentive compatible* if no student can possibly benefit by unilaterally misrepresenting her preferences (in the direct version of the mechanism), and *strategy-proof* if it is dominant-strategy incentive compatible. A *direct* revelation mechanism is one where each decision-maker is expected to report her preferences, while in an *indirect* mechanism decision-makers are asked to send messages other than directly their preferences.

The literature has explored the theoretical properties of the DA and TTC mechanisms in detail. In the school-allocation problem, in general, both are strategy-proof, that is submitting one's true preference order to the matchmaker is a dominant strategy for all students. The DA mechanism preserves stability at the cost of efficiency, while the TTC mechanism preserves efficiency at the cost of stability (Abdulkadiroğlu and Sönmez, 2003). For the preference profile in our experimen-

[18]By *student-stable* we refer to the stable matching preferred by students to all other stable matchings (Roth and Sotomayor, 1990). Our school-allocation problem has another stable matching, $\{(R_1, C), (R_2, A), (R_3, D), (H, B)\}$, which is preferred by schools to all other stable matchings. Given the one-sided definition for efficiency, the latter stable matching is not efficient.

[19]Section A in the online appendix describes the two algorithms in detail.

tal design, the DA and the TTC algorithms implement the same matching that is both stable and efficient.

The mechanisms based on the reverse algorithms (RDA and RTTC) induce students to submit the reverse of their preference order to the matchmaker. These equilibria are in dominant strategies, just like in the case of the DA and TTC mechanisms, and the equilibrium outcomes are identical to those implemented by the DA and TTC mechanisms. In other words, the DA and RDA mechanisms are identical in terms of efficiency and stability. So are the TTC and RTTC mechanisms. Moreover, all four mechanisms have equilibria in dominant strategies. Equilibrium behavior is the same in DA as in TTC, and in RDA as in RTTC.

Note that although matching theory typically assumes complete information, the fact that all the matching mechanisms considered here have equilibria in dominant strategies, means that for optimal behavior the decision-makers do not need any information about others' preferences, their objectives or rationality.

We summarize the theoretical predictions for our experimental treatments in two predictions.

Prediction 1. *Students submit their true preference ordering over schools in treatments based on the deferred-acceptance (DA) and top-trading-cycles (TTC) algorithms. They submit the reverse preference ordering in treatments based on the reverse deferred-acceptance (RDA) and reverse top-trading-cycles (TTC) algorithms.*

Prediction 2. *For the preference profile in our experimental design, all four analyzed matching mechanisms produce stable and efficient matchings as the final outcome.*

Note that the advantage of the experimental method is that we are not only able to test the above theoretical predictions, but we can also look behind the observed final decisions. Ultimately, our goal is to study strategy-proofness and the revelation principle in the experimental laboratory.

The theoretical literature leans towards mechanisms with equilibria in dominant strategies (Pathak, 2017), claiming that they constitute simple and practically non-strategic environments in which human decision-makers are very likely to behave in the predicted optimal way. The experimental literature seems to lend empirical support to these claims by reporting how often experimental subjects reported the assigned preference rankings truthfully.[20]

[20]For example, Chen and Sönmez (2006) test three well-known school-choice mechanisms (BOS,

Hypothesis 1. *In the school-allocation problem under the DA and the TTC mechanisms, students fully understand the strategic interaction and are able to find (and play) the equilibrium strategy.*

From a strictly theoretical point of view, the DA mechanism is not superior to the RDA mechanism, nor TTC is superior to RTTC. A cornerstone of mechanism design, the revelation principle, states that the equilibrium outcome of any indirect mechanism can be replicated by a direct mechanism in which truthful preference revelation is an equilibrium.[21] From a behavioral perspective, one might argue that certain mechanisms have easier or more intuitive equilibria than others, or that their equilibria coincide with natural focal points. However, we lack a clear definition of focal points and also criteria for ranking mechanisms according to their complexity.[22] We state the revelation principle for the here-analyzed mechanisms as a hypothesis to be tested in the experimental laboratory.

Hypothesis 2. Revelation principle. *The direct and reverse mechanisms (DA and RDA, and TTC and RTTC) implement the same matching in the school-allocation problem.*

3.2 Experimental findings

The experimental findings described in this section are based on formal two-sided statistical hypothesis tests: parametric t-tests for means, z-tests for proportions, and the non-parametric Kruskal-Wallis test. Note that our experimental design did not involve repetition or even interaction among participants, for that reason we do not control for time, learning and possible interdependence across observations.

If the related theoretical model is to be validated, experimental participants should report the assigned preference order (ABCD) to the DA and the TTC mechanisms, and its reverse (DCBA) to the RDA and RTTC mechanisms (prediction 1). Our data do not support these claims (table 2). Equilibrium play was observed in about two thirds of the cases under the DA mechanism which outperformed the other three mechanisms. TTC ranks second (46%), while RDA and RTTC share

DA and TTC) and, with the help of two treatments, show that in therms of inducing truth-telling in the laboratory, DA (72%, 56%) outperforms TTC (50%, 43%), which in turn performs better than BOS (14%, 28%).

[21] The revelation principle was introduced by Gibbard (1973) for mechanisms with equilibria in dominant strategies and later extended to Nash equilibria in Bayesian games (e.g., Myerson, 1981).

[22] Some might argue that mechanisms with equilibria in dominant strategies are "easier" to play than mechanisms that only have Nash equilibria, for instance. Recall, however, that our mechanisms are indistinguishable from each other in this respect.

the third position (with 16% and 26%, respectively). These findings are in line with the ones reported by Chen and Sönmez (2006) who found that DA induces more truthful (i.e., equilibrium) behavior than TTC. In summary, we clearly reject hypothesis 1 that the analyzed dominant-strategy incentive-compatible mechanisms induce equilibrium play. The following result offers a comparison among the four without theoretical reference purely based on experimental outcomes.

Result 1. *In terms of inducing equilibrium play, the DA mechanism is not worse than the TTC, which in turn is better than the RDA and the RTTC mechanisms which are undistinguishable from each other.*

	DA	TTC	RDA	RTTC
	68.00%	46.43%	16.33%	25.93%
DA	-	0.03	0.00	0.00
TTC	-	-	0.00	0.03
RDA	-	-	-	0.24

Table 2: Proportion of equilibrium play.
NOTE: p-values for pairwise two-sided comparisons of proportions reported under the percentages.

Comparing the "original" mechanisms to their reverse versions in terms of induced behavior reveals how much of the observed truthful "equilibrium behavior" under the DA and the TTC mechanisms is really due to rational strategizing assumed by theoretical models. We find a significant proportion of decisions-makers who report the assigned preference order to the matchmaker truthfully even under the reverse mechanisms where that is not optimal. Truth-telling is not only the modal strategy for both direct mechanisms, but it remains so even under RDA and ranks as the second most-frequently played strategy under RTTC (with a frequency of 22% closely behind the 26% of the equilibrium strategy). In other words, truth-telling is a strong focal point in the school-allocation problem.

One could argue that the reverse mechanisms pose a much larger challenge than the "original" ones. However, the proportion of participants who managed to solve the quiz correctly does not differ significantly between DA and RDA (p-value = 0.43) or between TTC and RTTC (p-value = 0.13). If anything, the observed proportion of correct answers is a slightly larger under the reverse mechanisms (table 3). As for participants' individual judgement in terms of the difficulty of the mechanisms, DA and RDA are not statistically different (p-value = 0.15), while TTC and RTTC are (p-value = 0.05). Interestingly, the latter comparison suggests that RTCC is more difficult.

	ORIGINAL	DA	TTC	REVERSE	RDA	RTTC
ABCD	TRUTH; EQ	68.00%	46.43%	TRUTH	30.61%	22.22%
BACD	DSB	22.00%	35.71%	NDSB	22.45%	14.81%
DCAB	-	-	-	DSB	16.33%	14.81%
DCBA	-	-	-	EQ	16.33%	25.93%
OTHER	-	10.00%	17.86%	-	14.28%	22.23%
QUIZ	-	64.00%	3.57%	-	71.43%	11.11%
DIFFICULT	-	4.36	5.14	-	5.22	6.33

Table 3: Declared preference orders (truth: ABCD).
NOTE: EQ: equilibrium; DSB: district-school bias; NDSB: naïve district-school bias; QUIZ: success rate in control quiz; DIFFICULT: answers (0-10) to "How difficult was the student allocation task?".

Result 2. *Approximately half of truth-telling observed in the deferred-acceptance and the top-trading-cycles mechanisms is the result of naïve (non-strategic) behavior.*

Once we separate the equilibrium strategy from the strong focal strategy of telling the truth, we can tell that optimal behavior (as predicted by theory) is rare. Overall, 21% of participants played the equilibrium strategy in the reverse mechanisms (16% in RDA, 26% in RTTC). The following result builds on this observation and on a 95% confidence interval around it.

Result 3. *Approximately 13-29% of participants act in sophisticated (strategic) manner.*

Answers to items in our post-experimental questionnaire sheds some more light on the reasoning behind the observed individual decisions. Recall that all four analyzed mechanisms have an equilibrium in dominant strategies. Those equilibria are "strong" in the sense that decision-makers can totally ignore the strategic interdependence of the situation: they do not need to know how others rank the possible outcomes, they do not even need to know whether the others are rational decisions-makers or not. One could argue that the situation that these mechanisms create are not (game-theoretic) games, but "simple" decision problems with a unique solution to them for each participant.

We wanted to know whether our experimental subjects understood that there exists a individual solution (or best strategy) to the school-allocation problem and that the solution does not depend on how the other decision-makers think about the problem and its outcomes. We asked participants to indicate on a scale from 0 to 10 how strongly they agree or disagree with the statement that "There was a best strategy for reporting preferences in the student allocation task." The extreme scores of 0 and 10 mean strong disagreement and strong agreement, respectively. The overall average of the collected scores shows the highest possible level of cluelessness as it lies at 4.83 and the 95% confidence interval around it includes score 5 (first row in table 4). Participants who ended up choosing the equilibrium strategy seem to be significantly more confident than those who did not, but even their score averages below 6.

Overall 61% of participants gave an affirmative answer to the question "Would you reconsider the way you acted if you knew each computer's most valued school?" (second row in table 4). According to the 95% confidence interval, more than half and as many as two thirds of decisions-makers would want to have more information about the others' preferences in a situation in which such information in totally unnecessary for making an optimal decision. Again, participants who submitted the equilibrium strategy seem to be significantly less likely to agree to the above question than those who submitted some other strategy. Nevertheless, almost half of them did say they would reconsider their decisions if they had more information.

As for reconsidering the chosen strategy, more than a fifth of our participants claimed they would do so if they had another chance (third row in table 4). Participants who played the equilibrium strategy were significantly more confident than the rest, but even in that group each tenth person said she would play differently next time.

	ALL		THEORY	EQUILIBRIUM PLAY		
	MEAN	(95% CONF.INT.)	REF.	YES		NO
STRATEGY	4.83	(4.36 ; 5.30)	10	5.83	$\neq_{0.00}$	4.18
INFORMATION	61.24%	(54.28% ; 67.89%)	0%	48.78%	$\neq_{0.00}$	69.29%
RECONSIDER	21.53%	(16.16% ; 27.73%)	0%	10.98%	$\neq_{0.00}$	28.35%
OBS.	209	-	-	82		127

Table 4: Answers to items in the questionnaire.

NOTE: STRATEGY: average agreement rate (0-10) to "There was a best strategy for reporting preferences in the student allocation task."; INFORMATION: % of YES answers to "Would you reconsider the way you acted if you knew each computer's most valued school?"; RECONSIDER: % of YES answers to "If you had another chance at the student allocation task would you act differently?". THEORY REF.: answers/values in line with game theory. $\neq_{p\text{-value}}$: statistically significant difference between groups; p-value for t-test for means (STRATEGY), z-test for proportions (INFORMATION, RECONSIDER).

Altogether merely 14 out of 209 participants (7%) played the equilibrium strategy and answered the three questions discussed above (and in table 4) in line with the assumption behind theoretical results. In conclusion, the following claim sharpens result 3 based on the 95% confidence interval around the point estimate.

Result 4. *Approximately 4-11% of participants act in sophisticated (strategic) manner and have full understanding of the "game".*

This means that the empirical success of direct mechanisms are in part due to naïve behavior, that is decision-makers who simply tell the truth instead of strategizing (that would also lead to the same conclusion).

	DA	TTC	RDA	RTTC
	72.00%	53.57%	55.10%	48.15%
DA	-	0.05	0.08	0.01
TTC	-	-	0.88	0.57
RDA	-	-	-	0.48

Table 5: Proportion of student-stable/efficient matchings.
NOTE: p-values for pairwise two-sided comparisons of proportions reported under the percentages.

We argue that the revelation principle does not hold as a general equivalence result between direct and non-direct mechanisms. Note that although the revelation principle (i.e., our hypothesis 2) is typically written in terms of outcome, it implicitly relies on fully rational decision-makers to be able to find the equilibrium strategy. In other words, the above experimental results raise serious doubts about its empirical relevance. As for a direct test, recall that, based on the usual assumption of decision-makers' rationality and sophistication, theory predicts that all four analyzed matching mechanisms produce stable and efficient matchings as the final outcome (prediction 2). Our experimental results clearly reject this prediction (table 5). The best-performing mechanism, DA, implemented stable and efficient matchings in about three quarters of the time, while the other three did so in about half of the time. Although comparisons based on statistical tests are not transitive in general, our data lead us conclude that, in terms of stability and efficiency, DA outperforms the other three mechanisms which are statistically identical. For the sake of precision, and because some might not consider a difference with a p-value of 0.08 statistically significant, we state these findings as follow.

Result 5. *In terms of stability and efficiency, the DA mechanism is not worse than the RDA, the TTC or the RTTC mechanisms which are undistinguishable from each other.*

It is worth noting that Chen and Sönmez (2006) find that the DA and TTC mechanisms perform equally well in terms of efficiency, and Calsamiglia et al. (2010) observe that DA produces stable matchings more often than TTC (in a school-allocation setting similar to ours in both studies). Although the conclusions based on comparisons among the numbers in table 5 are similar to those in the existing literature, our experimental design is not able to deliver sharp estimates for the stability (and efficiency) of the market outcome. Note that our experimental design targets individual behavior (i.e., strategy-proofness which is the main focus of our paper) instead of the aggregate outcome. It is built on a specific preference profile and only allows for untruthful behavior for the human decision-maker. With these features, in our school-choice problem there exist only two possible market outcomes: $\{(R_1, C), (R_2, B), (R_3, D), (H, A)\}$ and $\{(R_1, C), (R_2, A), (R_3, D), (H, B)\}$.[23] Thus, by design, not only just one market participant is allowed to strategize, but also the possible harm in terms of stability and efficiency caused by that strategizing behavior is bounded from above. Notwithstanding, the observed difference between the DA mechanism and its reverse version questions the empirical relevance of the revelation principle even if based on the observed outcomes we can not tell TTC and RTTC apart.

One could say that by definition (and design) strategy-proof direct mechanisms are robust to naïve behavior. Just as the popular direct mechanisms benefit from naïve truth-telling, (matching) mechanisms in general can benefit from various behavioral biases. Our reverse mechanisms, for instance, deliver the theoretical equilibrium outcome when the decision-maker suffers from naïve district-school bias, that is when she reports the induced preference order truthfully except that she moves her district school (at which she enjoys priority) to the top of the ranking. Note that in our school-allocation problem, with robots playing the equilibrium strategy, for the final outcome it only mattered how the human decision-maker ranked schools A and B with respect to each other. This is why hypothesis 2, the revelation principle written in terms of payoffs, is harder to refute by observing experimental outcomes. We argue, however, that the revelation principle is a theoretical result that builds on rationality, that is our hypothesis 1. By rejecting hypothesis 1, we also reject the revelation principle.

[23] Had our experimental design informed the human decision-maker about the entire preference profile, instead of keeping the robots' preferences hidden, our experimental subjects would have faced a situation in which multiple weakly dominant strategies exist. Essentially, all strategies (for the human decision-maker) that rank A above B in the direct mechanisms (and those that rank B above A in the reverse mechanisms) would be weakly dominant in such a game.

4 Discussion and concluding remarks

The theory of market design relies on the rationality of decision-makers who are expected to be able to fully understand the decision problem at hand and find the dominant strategy before playing it. We have argued that the *usual* experimental tests of predictions derived from this theory test joint hypotheses and, for that reason, confuse observed choices for well-reasoned optimal behavior. While strategy-proof mechanisms are preferred by theorists for having a strong and supposedly easy equilibrium (e.g., Vickrey, 1961) and more generally for not giving more and better chances to sophisticated people than to naïve decision-makers (e.g., Pathak and Sönmez, 2008; Pathak, 2017), these claims are based on assumptions that are only valid for a very small fraction of decision-makers.

Our findings from the experimental laboratory contribute to the debate on the extent of preference manipulation in school-choice mechanism.[24] That sophisticated and unsophisticated decision-makers coexist does not come as a surprise (e.g., Abdulkadiroğlu et al., 2006). Without some reliable estimates for the proportions in which they are present in observational data, however, it is impossible to perform reliable normative welfare analysis on school-allocation systems.

With the help of controlled laboratory experiments, we have shown that roughly half of truth-telling (in the typical school-choice problem based on DA or TTC) seems to be the result of naïve (non-strategic) behavior and that only a small fraction (around 4-11%) of decision-makers act in sophisticated manner and have full understanding of the strategic properties of the situation. These results question the applicability and relevance of matching theory as a whole. In other words, although some may argue that strategy-proof mechanisms offer a reasonable solution by inviting both sophisticated and naïve decision-makers to act in line with its equilibrium, it is not clear whether they would perform better than other (much simpler) mechanisms in real life. After all, a large part of the population belongs to other groups in terms of strategic sophistication (between or beside the above-mentioned extremes), and mechanisms designed for them could outperform the "standard" strategy-proof mechanism.[25] In summary, our experimental results question the

[24]For example, the analysis of the consequences of changes in Chicago's assignment system in 2009, those of Barcelona's and Beijing's adoption of the Boston mechanism, reviewed by Pathak (2017), raises important questions related to the very foundations of mechanism design without being able to deliver precise estimates on the proportion of sophisticated decision-makers that exist in the population.

[25]Clearly, successful mechanisms should tolerate certain behavioral faults (McFadden, 2009), however with the help of the very same experimental design, we are unable to establish a new set of

empirical relevance of strategy-proofness and the revelation principle.

Acknowledgements

Financial support from the Waseda Institute of Political Economy and the JSPS KAKENHI Grant No. 17K03634 has made our experimental sessions possible.

References

[1] Abdulkadiroğlu, A., Pathak, P.A., Roth, A.E. & Sönmez, T., 2006. Changing the Boston school choice mechanism, *NBER Working Paper No. 11965.*

[2] Abdulkadiroğlu, A. & Sönmez, T., 2003. School choice: A mechanism design approach. *American Economic Review*, 93(3): 729-747.

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

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

[5] Chen, Y., Liang, Y. & Sönmez, T., 2016. School choice under complete information: An experimental study. *Journal of Mechanism and Institution Design*, 1(1): 45-82.

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

[7] Fischbacher, U., 2007. z-Tree - Zurich toolbox for readymade economic experiments - Experimenter's manual. *Experimental Economics*, 10(2): 171-178.

[8] Gale, D. & Shapley, L., 1962. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1): 9-15.

[9] Gibbard, A., 1973. Manipulation of voting schemes: a general result. *Econometrica*, 41(4): 587-601.

empirically justified assumptions. Obviously strategy-proof mechanisms (Li, 2017) do not seem to help as stable matching mechanisms have been shown not to be obviously strategy-proof (Ashlagi and Gonczarowski, 2018).

[10] Guillen, P. & Hakimov, R., 2017. Not quite the best response: truth-telling, strategy-proof matching, and the manipulation of others. *Experimental Econconomics*, 20: 670-686.

[11] Guillen, P. & Hakimov, R., 2018. The effectiveness of top-down advice in strategy-proof mechanisms: A field experiment. *European Economic Review*, 101: 505-511.

[12] Guillen, P. & Hing, A., 2014. Lying through their teeth: Third party advice and truth telling in a strategy proof mechanism. *European Economic Review*, 70(C): 178-185.

[13] Hakimov, R. & Kübler, D., 2019. Experiments on matching markets: A survey. *mimeo*

[14] Hassidim, A., Romm, A. & Shorrer, R.I., 2018. 'Strategic' behavior in a strategy-proof environment. *Available at SSRN: http://dx.doi.org/10.2139/ssrn.2784659*

[15] Hassidim, A., Marciano, D., Romm, A. & Shorrer, R.I., 2017. Mistakes in dominant-strategy mechanisms. The mechanism is truthful, why aren't you? *American Economic Review: Papers & Proceedings*, 107(5): 220-224.

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

[17] McFadden, D., 2009. The human side of mechanism design: a tribute to Leo Hurwicz and Jean-Jacque Laffont. *Review of Economic Design*, 13(1-2): 77-100.

[18] Myerson, R., 1981. Optimal auction design, *Mathematics of Operations Research*, 6(1): 58-73.

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

[20] Pais, J., Pintér, Á. & Veszteg, R.F., 2011. College admissions and the role of information: An experimental study. *International Economic Review*, 52(3): 713-737.

[21] Pathak, P.A., 2017. What really matters in designing school choice mechanisms. In B. Honoré, Pakes, A., Piazzesi, M. & Samuelson, L. (Eds.), *Advances in Economics and Econometrics: Eleventh World Congress* (Econometric Society Monographs, 176-214), Cambridge University Press.

[22] Pathak, P.A. & Sönmez, T., 2008. Leveling the playing field: Sincere and sophisticated players in the Boston mechanism. *American Economic Review*, 98(4): 1636-1652.

[23] Pathak, P.A. & Sönmez, T., 2013. School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation. *American Economic Review*, 103(1): 80-106.

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

[25] 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.

[26] Roth, A.E. & Sotomayor, M.A.O., 1990. *Two-sided matching, A study in game-theoretic modeling and analysis*, Cambridge University Press.

[27] Roth, A.E. 1991. A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the United Kingdom. *American Economic Review*, 81(3): 415-440.

[28] Roth, A.E., 2015. Experiments in market design. In Kagel, J.H. & Roth, A.E. (Eds.), *The handbook of experimental economics* (Volume 2., 290-347), Princeton University Press.

[29] Shapley, L. & Scarf, H., 1974. On cores and indivisibility. *Journal of Mathematical Economics*, 1(1): 23-37.

[30] Shorrer, R.I. & Sóvágó, S. 2018. Obvious mistakes in a strategically simple college admissions environment: Causes and consequences. *Available at SSRN: http://dx.doi.org/10.2139/ssrn.2993538*

[31] Vickrey, W., 1961. Counterspeculation, auctions, and competitive sealed tenders, *Journal of Finance*, 16(1): 8-37.

A Matching algorithms

This section describes the deferred-acceptance and the top-trading-cycles algorithms with the help of the text used in the experimental instructions.[26]

A.1 Deferred acceptance (DA)

- In this experiment, students in each group belong to a specific school districts. Your "local school" and those of the computer players are going to be indicated on the computer screen.
  NOTE: In this simulation, each school has one slot, but it can happen that more than one student lives in the same district.
- In each group, every student submits a ranking of schools which is used in the following procedure to allocate school places.
- The priority order for each school is separately determined as follows:
  - High Priority Level: Participants who live within the school district.
  - Low Priority Level: Participants who do not live within the school district. The priority among the Low priority students is based on their respective order in a fair lottery. This means each participant has an equal chance of being the first in the line, the second in the line, ..., as well as the last in the line. The computer is going to determine the outcome of this fair lottery. You will be informed about the outcome of the lottery on the screen.
- Once the priorities are determined, the allocation of school slots is obtained as follows:
  - An application to the highest ranked school on the submitted ranking is sent for each participant.
  - Through out the allocation process, a school can hold no more applications than its number of slots. If a school receives more applications than its capacity, then it rejects the students with lowest priority orders. The remaining applications are retained.
  - Whenever an applicant is rejected at a school, his application is sent to the next highest school on his/her submitted ranking.

[26]Note that we closely follow the set of instructions used by Chen et al. (2016).

- Whenever a school receives new applications, these applications are considered together with the retained applications for that school. Among the retained and new applications, the lowest priority ones in excess of the number of the slots are rejected, while remaining applications are retained.
- The allocation is finalized when no more applications can be rejected. Each participant is assigned a slot at the school that holds his/her application at the end of the process.

A.2 Top trading cycles (TTC)

- In this experiment, students in each group belong to a specific school districts. Your "local school" and those of the computer players are going to be indicated on the computer screen.
  NOTE: In this simulation, each school has one slot, but it can happen that more than one student lives in the same district.
- In each group, every student submits a ranking of schools which is used in the following procedure to allocate school places.
- Each student is first tentatively assigned to the school within his/her respective district.

  Next, the submitted rankings are used to determine the school allocation through exchanges. The order in which these exchanges are considered is determined by a fair lottery. This means each student has an equal chance of being the first in the line, the second in the line, ..., as well as the last in the line. You are going to receive information about the outcome of the lottery on the screen.
- The specific allocation process is explained below.
  - Initially all slots are available for allocation.
  - All students are ordered in a queue based on the order in the lottery.
  - Next, an application to the highest ranked school in the submitted ranking is submitted for the student at the top of the queue.
    * If the application is submitted to his district school, then his tentative assignment is finalized (thus he is assigned a slot at his district

school). The student and his assignment are removed from subsequent allocations. The process continues with the next student in line.

* If the application is submitted to another school, say school B, and that school has a vacant slot (it has no other student tentatively assigned), then the requester is assigned to school B. The student and his assignment are removed from subsequent allocations. The process continues with the next student in line.
* If the application is submitted to another school, say school C, and that school has no vacant slot (it has another student tentatively assigned), then the first student in the queue who tentatively holds a slot at School C is moved to the top of the queue directly in front of the requester.

– Whenever the queue is modified, the process continues similarly: An application is submitted to the highest ranked school with available slots for the student at the top of the queue.

  * If the application is submitted to his district school, then his tentative assignment is finalized. The process continues with the next student in line.
  * If the application is submitted to another school, say school B, and that school has a vacant slot (it has no other student tentatively assigned), then the requester is assigned to school B. The student and his assignment are removed from subsequent allocations. The process continues with the next student in line.
  * If the application is submitted to another school, say school C, and that school has no vacant slot (it has another student tentatively assigned), then the first student in the queue who tentatively holds a slot at school C is moved to the top of the queue directly in front of the requester. This way, each student is guaranteed an assignment based on the preferences indicated in the submitted ranking.

– An exchange is obtained when a cycle of applications are made in sequence, e.g., I apply to John‘s district school, John applies to your district school, and you apply to my district school. In this case, the exchange is completed and the students as well as their assignments are removed from subsequent allocations.

- The process continues until all students are assigned a school slot.

B Instructions[27]

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 addition to a ¥700 participation fee). 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

- Please the read these instructions first before proceeding to the computer program. You may consult the instructions at any time during the experiment.
- Each participant is randomly assigned to a group of 4 decision-makers. Each group includes 3 computers and a participant. This means that you are going to interact with 3 computers in this experiment. The computers will act to maximize their expected gain.
- In this simulation, each member of your group acts as a student who is looking for a slot at one of the schools. 4 school slots are available across 4 schools (A, B, C, D) in each group. There is one slot available at each school.
- You will receive a cash payment at the end of the experiment. Your payoff amount depends on the school slot you hold at the end of the experiment. Payoff amounts are going to be outlined on the computer screen.

  NOTE: Different participants and different computers may or may not have different payoff tables. That is, payoff by school may or may not be different for different participants and different computers.
- During the experiment, each participant first completes a quiz which consists of an allocation instance for you to solve. You will have up to 10 minutes to solve it, and will receive ¥100 if you submit the correct answer.

[27] The following text was used to create the Japanese version of the experimental instructions for our RDA (reverse deferred-acceptance algorithm) treatment. The documents that we prepared for our treatments are available upon request both in English and Japanese. We closely follow the set of instructions used by Chen et al. (2016) in all of them.

- After you have finished the quiz you may start the school allocation task. You will have up to 20 minutes to complete the school allocation task.

- After all participants have completed the school allocation task, the experimenter starts the allocation process.

- Once the allocations are determined, the experimenter informs each participants of his/her allocation slot and respective payoff. The school allocation will be performed based on the following school allocation method.

Allocation Method

- In this experiment, students in each group belong to a specific school districts. Your "local school" and those of the computer players are going to be indicated on the computer screen.

  NOTE: In this simulation, each school has one slot, but it can happen that more than one student lives in the same district.

- In each group, every student submits a ranking of schools which is used in the following procedure to allocate school places.

- Each student is first tentatively assigned to the school within his/her respective district.

  Next, the submitted rankings are used to determine the school allocation through exchanges. The order in which these exchanges are considered is determined by a fair lottery. This means each student has an equal chance of being the first in the line, the second in the line, ..., as well as the last in the line. You are going to receive information about the outcome of the lottery on the screen.

- The specific allocation process is explained below.

  - Initially all slots are available for allocation.
  - All students are ordered in a queue based on the order in the lottery.
  - Next, an application to the lowest ranked school in the submitted ranking is submitted for the student at the top of the queue.
    - If the application is submitted to his district school, then his tentative assignment is finalized (thus he is assigned a slot at his district

school). The student and his assignment are removed from subsequent allocations. The process continues with the next student in line.

* If the application is submitted to another school, say school B, and that school has a vacant slot (it has no other student tentatively assigned), then the requester is assigned to school B. The student and his assignment are removed from subsequent allocations. The process continues with the next student in line.
* If the application is submitted to another school, say school C, and that school has no vacant slot (it has another student tentatively assigned), then the first student in the queue who tentatively holds a slot at School C is moved to the top of the queue directly in front of the requester.

- Whenever the queue is modified, the process continues similarly: An application is submitted to the lowest ranked school with available slots for the student at the top of the queue.

  * If the application is submitted to his district school, then his tentative assignment is finalized. The process continues with the next student in line.
  * If the application is submitted to another school, say school B, and that school has a vacant slot (it has no other student tentatively assigned), then the requester is assigned to school B. The student and his assignment are removed from subsequent allocations. The process continues with the next student in line.
  * If the application is submitted to another school, say school C, and that school has no vacant slot (it has another student tentatively assigned), then the first student in the queue who tentatively holds a slot at school C is moved to the top of the queue directly in front of the requester. This way, each student is guaranteed an assignment based on the preferences indicated in the submitted ranking.

- An exchange is obtained when a cycle of applications are made in sequence, e.g., I apply to John's district school, John applies to your district school, and you apply to my district school. In this case, the exchange is completed and the students as well as their assignments are removed from subsequent allocations.

– The process continues until all students are assigned a school slot.

An Example

We go through a simple example to illustrate how the allocation method works.

Students and Schools: In this example, there are six students, 1-6, and four schools, Clair, Erie, Huron and Ontario.

Student ID Number: 1, 2, 3, 4, 5, 6	Schools: Clair, Erie, Huron, Ontario

Slots and Residents: There are two slots each at Clair and Erie, and one slot each at Huron and Ontario. Residents of districts are indicated in the table below.

School	Slot 1	Slot 2	District Residents
Clair	□	□	1 2
Erie	□	□	3 4
Huron	□		5
Ontario	□		6

Tentative assignments: Students are tentatively assigned slots at their district schools.

School	Slot 1	Slot 2	
Clair	1	2	Students 1 and 2 are **tentatively assigned** a slot at Clair;
Erie	3	4	Students 3 and 4 are **tentatively assigned** a slot at Erie;
Huron	5	-	Student 5 is **tentatively assigned** a slot at Huron;
Ontario	6	-	Students 6 is **tentatively assigned** a slot at Ontario.

Lottery: The lottery produces the following order.

1 - 2 - 3 - 4 - 5 - 6

Submitted School Rankings: The students submit the following school rankings:

	Last Choice	3rd Choice	2nd Choice	1st Choice
Student 1	Huron	Clair	Ontario	Erie
Student 2	Huron	Ontario	Clair	Erie
Student 3	Ontario	Clair	Erie	Huron
Student 4	Huron	Clair	Ontario	Erie
Student 5	Ontario	Huron	Clair	Erie
Student 6	Clair	Erie	Ontario	Huron

This allocation method consists of the following steps:

Step 1: A fair lottery determines the following student order: 1-2-3-4-5-6. Student 1 has ranked Huron as his last choice. However, the only slot at Huron is tentatively held by student 5. So student 5 is moved to the top of the queue.

Step 2: The modified queue is now 5-1-2-3-4-6. Student 5 has ranked Ontario as his last choice. However, the only slot at Ontario is tentatively held by student 6. So student 6 is moved to the top of the queue.

Step 3: The modified queue is now 6-5-1-2-3-4. Student 6 has ranked Clair as her last choice. The two slots at Clair are tentatively held by students 1 and 2. Between the two, student 1 is ahead in the queue. So student 1 is moved to the top of the queue.

Step 4: The modified queue is now 1-6-5-2-3-4. Remember that student 1 has ranked Huron as his last choice. A cycle of applications is now made in sequence in the last three steps: student 1 applied to the tentative assignment of student 5, student 5 applied to the tentative assignment of student 6, and student 6 applied to the tentative assignment of student 1. These exchanges are carried out: student 1 is assigned a slot at Huron, student 5 is assigned a slot at Ontario, and student 6 is assigned a slot at Clair. These students as well as their assignments are removed from the system.

Step 5: The modified queue is now 2-3-4. There is one slot left at Clair and two slots left at Erie. Student 2 applies to Clair, which is her last choice between the two schools with remaining slots. Since student 2 tentatively holds a slot at Clair, her tentative assignment is finalized. Student 2 and her assignment are removed from the system.

Step 6: The modified queue is now 3-4. There are two slots left at Erie. Student 3 applies to Erie, which is the only school with available slots. Since Student 3 tentatively holds a slot at Erie, her tentative assignment is finalized. Student 3 and her assignment are removed from the system.

Step 7: The only remaining student is student 4. There is one slot left at Erie. Student 4 applies to Erie for the last available slot. Since Student 4 tentatively holds a slot at Erie, his tentative assignment is finalized. Student 4 and his assignment are removed from the system.

Final assignment: Based on this method, the final allocations are:

Student	1	2	3	4	5	6
School	Huron	Clair	Erie	Erie	Ontario	Clair

Illustration

	Queue	Available Slots	The top student in the queue applies to a school.	At the end of the step
Step 1	1-2-3-4-5-6	Clair Clair Erie Erie Huron Ontario	1 applies to her last choice Huron, which is tentatively assigned to 5.	5 comes to the top. 1 2 3 4 **5** 6
Step 2	5-1-2-3-4-6	Clair Clair Erie Erie Huron Ontario	5 applies to her last choice Ontario, which is tentatively assigned to 6.	6 comes to the top. 5 1 2 3 4 **6**
Step 3	6-5-1-2-3-4	Clair Clair Erie Erie Huron Ontario	6 applies to her last choice Clair, which is tentatively assigned to 1 and 2.	1 comes to the top. 6 5 **1** 2 3 4
Step 4	1-6-5-2-3-4	Clair Clair Erie Erie Huron Ontario	A cycle happens in the last 3 steps.	1 gets a slot at Huron. 5 gets a slot at Ontario. 6 gets a slot at Clair.
Step 5	2-3-4	Clair Erie Erie	2 applies to her 2nd choice Clair, because her last and 3rd choices (Huron and Ontario) are no longer available.	2 gets a slot at Clair, because she is a resident in Clair.
Step 6	3-4	Erie Erie	3 applies to Erie which is still available.	3 gets a slot at Erie, because he is a resident in Erie.
Step 7	4	Erie	4 applies to Erie.	4 gets a slot at Erie, because she is a resident in Erie.

Final assignment: Based on this method, the final allocations are:

Student	1	2	3	4	5	6
School	Huron	Clair	Erie	Erie	Ontario	Clair

Quiz

Please find the correct allocation for the instance explained below. You will earn ¥100 for a correct answer.

There are 6 students (ID numbers from 1 to 6), and 3 schools (school A, school B and school C) with two places each. Students 2 and 3 live in the district of School A, students 4 and 5 live in the district of School B and, finally, students 1 and 6 live in the district of School C.

School	District Residents	
A	2	3
B	4	5
C	1	6

The lottery determined the following order (student IDs): 5 - 6 - 2 - 1 - 3 - 4.

Each student submitted a school ranking. These are given on the **Quiz** page on your computer screen.

You have up to 10 minutes to determine the final allocation. If you have any questions raise your hand and we will come to you. However, the experimenter will not assist you with the task.

After completing this quiz you will have up to 20 minutes to complete the main school allocation task. Complete the task at your own pace.

After the allocation task there will be a short questionnaire. When you have completed the questionnaire raise your hand and the experimenter will come over to conclude the experiment.