Improving Efficiency of On-Campus Housing: An Experimental Study

By YAN CHEN AND TAYFUN SÖNMEZ*

This paper investigates a class of matching problems—the assignment of indivisible items to agents where some agents have prior claims to some of the items. As a running example, we will refer to the indivisible items as houses. House allocation problems are not only of theoretical interest, but also of practical importance. A house allocation mechanism assigns a set of houses (or offices, tasks, etc.) to prospective tenants, allotting at most one house to each tenant. Rents are exogenously given and there is no medium of exchange, such as money. In general some houses will have existing tenants, some houses will be empty, and some applicants for housing will be new (e.g., freshmen). The canonical examples are assignment of college students to dormitory rooms and public housing units. Other examples are assignment of offices and tasks to individuals.

Many universities in the United States employ some variant of a mechanism called the *random serial dictatorship with squatting rights* (RSD) to allocate dormitory rooms. Each existing tenant can either keep her house or enter the applicant pool. Each applicant is randomly given a (possibly seniority-weighted or GPA-weighted) priority and each is assigned, in priority order, her top choice among the houses that remain. This mechanism is strategy-proof (i.e., dominant strategy incentive compatible): truthful preference revelation is a dominant strategy for each applicant (Lin Zhou, 1990). On the negative side, a tenant who enters the lottery may end up with a house that is worse than her current house. As a result not every existing tenant joins the applicant pool, potential gains from trade are lost, and the mechanism yields Pareto-inefficient outcomes (Atila Abdulkadiroğlu and Sönmez, 1999). While we are not aware of any systematic field studies of this mechanism, there is some evidence for this inefficiency. For example, at the University of Michigan, where RSD is used to assign undergraduate students to residence halls, there is a restricted after-market for leases.[1]

Motivated by these observations, Abdulkadiroğlu and Sönmez (1999) propose a simple mechanism, the top trading cycles (TTC) mechanism, as a superior alternative. In this mechanism, applicants are again prioritized and are given their top choice in priority order. This process continues until someone requests an existing tenant's house. In this case, the existing tenant is moved to the top of the priority queue, directly in front of the requester. If a cycle of requests is formed (e.g., I want John's house, John wants your house, and you want my house), all members of the cycle are given what they want, and their new houses are removed from the system. This mechanism is theoretically superior to the former: it is strategy-proof, individually rational, and Pareto efficient.

* Chen: School of Information, University of Michigan, 550 East University Avenue, Ann Arbor, MI 48109 (e-mail: yanchen@umich.edu); Sönmez: Department of Economics, Koç University, Sarıyer, 80910, Istanbul, Turkey (e-mail: tsonmez@ku.edu.tr). We thank Rachel Croson, Julie Cullen, Catherine Eckel, Elisabeth Hoffman, Lutz Kilian, David Lucking-Reiley, David Pennock, Laura Razzolini, Stephane Robin, Larry Samuelson, Sara Solnick, Lise Vesterlund, two anonymous referees, anonymous referees of the National Science Foundation, and seminar participants at the University of Aarhus, University of Copenhagen, Harvard University, Hong Kong University of Science and Technology, Lund University, University of Michigan, Rice University, University of St. Gallen, University of Zürich, the 1999 and 2000 ESA meetings, and SED 2000 meetings for helpful comments and discussions. Yuri Khoroshilov, Emre Sucu, Chun-Chih Chen, Sharon Lam, and Asad Qayyum provided excellent research assistance. Financial support from the National Science Foundation through Grant No. SBR-9805586 to Chen and SES-9904214 to Sönmez are gratefully acknowledged. Any remaining errors are our own.

[1] The Michigan *Reapplication Lease Renewal Program* (University Housing, 1999, p. 4) states that a returning student can only reassign a lease to a newly entering student. Returning students are not allowed to buy or sell leases from each other. In 1999 the application and assignment process was in February and March, while the after-market did not begin until August 2.

In theory the TTC mechanism performs better than RSD. However, it is not clear whether the TTC mechanism remains superior in practice when both mechanisms are implemented among boundedly rational agents. There have been many examples where a dominant strategy mechanism does not perform well in the laboratory because subjects do not use their dominant strategies. Well-known examples include the sealed-bid second-price auctions (John H. Kagel, 1995) and the pivotal mechanism (Chen, 2002).[2] Before offering the mechanism for actual use, we would like to observe and evaluate its performance in the context of actual decision problems faced by real people with real incentives. This motivated the research reported in this paper, where a laboratory experiment is designed both to test the theory and to address a class of interesting real-world problems. In Alvin E. Roth's (1995) words, these experiments are designed both to "speak to theorists" and to "whisper in the ears of Princes."

The house allocation problem belongs to the general class of matching problems. While we are not aware of any previous experimental studies of the house allocation problems, there have been experimental studies of other one-sided and two-sided matching problems motivated by various real-world applications. Mark Olson and David Porter (1994) compare the following four mechanisms in the context of the one-sided matching problem motivated by the Jet Propulsion Laboratory's problem of assigning time slots on its Deep Space Network of large antennas: generalized versions of the Vickrey and the English auctions, the serial dictatorship, and the chit mechanism. They find that the generalized auctions result in efficient allocations at the expense of consumer surplus, while the serial dictatorship and the chit mechanism are significantly less efficient in their assignment. Note that the serial dictatorship studied in Olson and Porter (1994) differs from our RSD because it does not deal with existing tenants with squatting rights.

We are aware of five experimental studies of two-sided matching problems. Haig R. Nalbantian and Andrew Schotter (1995) investigate the following three market-like mechanisms, each of which may be used to match teams and professional baseball players in their free-agent year: the current free-agency system, a complete-information English auction, and the Herman B. Leonard (1983)-Gabrielle Demange and David Gale (1985) mechanism. They show that the efficiency differences of these mechanisms are not statistically significant. The current free-agency system is successful in avoiding no-match outcomes, but it is less successful in matching people in an optimal manner. The opposite is true of the Leonard-Demange-Gale mechanism and the performance of the English auction is in between. Glenn W. Harrison and Kevin A. McCabe (1996) investigate the student-optimal stable mechanism (Gale and Lloyd S. Shapley, 1962) and show that profitable misrepresentation of preferences becomes difficult as markets get larger. Motivated by the field experiments matching new physicians to hospitals in the United States (Roth, 1984) and in the United Kingdom (Roth, 1991), Kagel and Roth (2000) compare the following two mechanisms: a priority-matching mechanism (of the kind unsuccessfully used in Newcastle) and a stable matching mechanism (of the kind successfully used in Edinburgh). In addition to reproducing the field observations, their experimental observations also provide new insights to the transition phase not observed in field experiments. M. Utku Ünver (2001) investigates the two mechanisms analyzed by Kagel and Roth (2000) as well as an unstable linear-programming mechanism (of the kind successfully used in London). In his experiment the unstable (and yet successful) London linear-programming mechanism performs no better than the unstable (and unsuccessful) Newcastle priority-matching mechanism in preventing early contracts although it performs better in terms of overall efficiency. Finally Ernan Haruvy et al. (2001) investigate the entry-level labor market for American federal law clerks and show that a centralized stable matching mechanism may not be as successful in preventing early contracts as the current decentralized system due to market-specific coordination issues.

Our experiment differs from the above six experiments in both the particular mechanisms

[2] Whether subjects use their dominant strategies is closely related to the transparency of the dominant strategy and the complexity of the mechanism, which are largely empirical questions.

studied, and the potential field applications of the mechanisms.

The paper is organized as follows. Section I reviews the theoretical properties of the two mechanisms. Section II contains a description of the experimental design. Section III summarizes the main results of the experiments. Section IV concludes the paper.

I. Theoretical Properties of the Mechanisms

We first introduce the model and then describe two mechanisms in this context.

A *house allocation problem with existing tenants* (Abdulkadiroğlu and Sönmez, 1999) consists of:

1. a finite set of existing tenants I_E,
2. a finite set of new applicants I_N,
3. a finite set of occupied houses $H_O = \{h_i\}_{i \in I_E}$,
4. a finite set of vacant houses H_V, and
5. a list of strict preference relations $P = (P_i)_{i \in I_E \cup I_N}$.

Here $h_i \in H_O$ is the house that is currently occupied by the existing tenant $i \in I_E$ who is entitled to keep it. In other words, the existing tenant i has the squatter's rights over house h_i. Each agent has use for one and only one house and has strict preferences over all houses. Without loss of generality we assume that the number of agents is the same as the number of houses. A *house allocation* is an assignment of houses to agents such that each agent is assigned a distinct house.

Two special cases of this model deserve mentioning. A *housing market* (Shapley and Herbert E. Scarf, 1974) is a special case where there are only existing tenants and occupied houses. A *house allocation problem* (Aanund Hylland and Richard Zeckhauser, 1979) is the other extreme case where there are only new applicants and vacant houses.

It is the responsibility of a centralized clearing house, e.g., the housing office, to allocate the houses among the agents. A house allocation mechanism consists of a strategy space for each agent and an outcome function which selects a lottery over house allocations for each strategy-tuple. In a direct mechanism agents simply state their preferences over houses. A house allocation mechanism is (*ex post*) *Pareto efficient* if it gives positive weight to only Pareto-efficient house allocations. A house allocation mechanism is (*ex post*) *individually rational* if it assures every existing tenant a house that is at least as good as her own. A direct house allocation mechanism is *strategy-proof* (or dominant strategy incentive compatible) if no agent can ever benefit by misrepresenting her preferences.

A. *Random Serial Dictatorship with Squatting Rights*

The *random serial dictatorship with squatting rights* (RSD) mechanism is commonly used in real-life applications of house allocation problems with existing tenants. Some examples include undergraduate housing at Carnegie Mellon University, Duke University, the University of Michigan, Northwestern University, and the University of Pennsylvania. The mechanism involves four steps:

1. Each existing tenant decides whether she will enter the housing lottery or keep her current house. Those who prefer keeping their houses are assigned their houses. All other houses become available for allocation.
2. Every agent who enters the lottery reports her preferences over all available houses.
3. An ordering of all agents who enter the lottery is randomly chosen from a given distribution of orderings. This distribution may be uniform or it may favor some groups (such as seniors) over others.
4. Available houses are allocated to agents based on their reported preferences and the chosen ordering: The first agent is assigned her top choice, the next agent is assigned her top choice among the remaining houses, and so on.

While this mechanism is very popular in real-life applications, it suffers from a major deficiency. Since it does not guarantee each existing tenant a house that is at least as good as her own, some tenants may choose to keep their houses even though they wish to move, and this may result in loss of potentially large gains from trade. Hence the RSD mechanism is neither individually rational nor Pareto efficient. It is, however, strategy-proof: truthful preference revelation is a dominant strategy for every agent who enters the lottery.

B. *Top Trading Cycles Mechanism*

The inefficiency of the RSD is caused by the mechanism's failure to guarantee each existing tenant a house at least as good as the one she already holds. To achieve efficiency we must address this "deficiency." This is the motivation for the *top trading cycles* (TTC) mechanism (Abdulkadiroğlu and Sönmez, 1999):

1. Each agent reports her preferences over all houses.
2. An ordering of agents is randomly chosen from a given distribution of orderings.
3. For any given preference list and ordering, the outcome is obtained using the following *you request my house–I get your turn* algorithm:

    (a) Assign the first agent her top choice, the second agent her top choice among the remaining houses, and so on, *until someone requests the house of an existing tenant.*
    (b) If at that point the existing tenant whose house is requested is already assigned another house, then do not disturb the procedure. Otherwise modify the remainder of the ordering by inserting the existing tenant to the top of the line and proceed with the procedure.
    (c) Similarly, insert any existing tenant who is not already served at the top of the line once her house is requested.
    (d) If at any point a loop forms, it is formed by exclusively existing tenants and each of them requests the house of the tenant who is next in the loop. (A *loop* is an ordered list of agents $(i_1, i_2, \ldots, i_k)$ where agent i_1 requests the house of agent i_2, agent i_2 requests the house of agent i_3, ... , agent i_k requests the house of agent i_1.) In such cases remove all agents in the loop by assigning them the houses they request and proceed with the procedure.

The key innovation in this mechanism is that an existing tenant whose current house is requested is upgraded to the first place at the remaining of the line before her house is allocated. As a result the TTC mechanism is individually rational, as it assures every existing tenant a house that is at least as good as her own.[3]

The TTC mechanism is also strategy-proof[4] and Pareto efficient. Another desirable feature is that it coincides with the competitive mechanism when restricted to housing markets[5] and with random serial dictatorship when restricted to house allocation problems.[6]

In the experiment we use a variant of the TTC mechanism which gives the existing tenants the option to opt out and keep their current houses (as in RSD). Since opting out is a weakly dominated strategy for the existing tenant, theoretically this variant should yield the same outcome. There are three reasons for using this variant. First, it keeps the size of the strategy space in the two mechanisms the same, which

[3] Alternatively one can fix the inefficiency caused by the RSD mechanism using the following individually rational mechanism: First construct an initial allocation by assigning each existing tenant her own house and randomly assigning a vacant house to each newcomer and next choose the unique competitive allocation of the induced housing market as the final outcome. Sönmez and Ünver (2001) show that this mechanism is equivalent to a special case of the TTC mechanism.

[4] Indeed the TTC mechanism cannot be manipulated by a group of agents either: When a group of agents jointly misrepresent their preferences, if any of the agents profit then at least one of them strictly loses.

[5] In a housing market there are only existing tenants and occupied houses. The *competitive allocation* is unique for each housing market and it coincides with the unique core allocation (Roth and Andrew Postlewaite, 1977). Moreover the *competitive mechanism* is strategy-proof (Roth, 1982) and the only one which is also Pareto efficient and individually rational (Jinpeng Ma, 1994). As a result, the competitive mechanism is considered the key mechanism for housing markets.

[6] In a house allocation problem there are only new applicants and vacant houses. The key mechanism in this context is the *random serial dictatorship:* Randomly order all agents and assign the first agent her top choice, the next agent her top choice among the remaining houses, and so on. Note that squatting rights do not apply in this context. This procedure is equivalent to randomly choosing an initial allocation and choosing the competitive allocation of the induced housing market (Abdulkadiroğlu and Sönmez, 1998). The random serial dictatorship always selects Pareto-efficient allocations and it is strategy-proof. Moreover its various versions are commonly used in real-life applications of house allocation problems.

TABLE 1—PAYOFF TABLE FOR ALL AGENTS—ORIGINAL ENVIRONMENT

Type of houses:		A	B	C	D	E	F	G	H
Existing tenants	#1	[6]	3	8	9	15	5	1	12
	#2	6	[5]	9	8	3	12	15	1
	#3	1	3	[9]	15	5	6	12	8
	#4	5	9	15	[12]	3	6	8	1
	#5	5	1	12	9	[3]	15	8	6
	#6	15	6	9	8	1	[12]	3	5
	#7	3	5	6	12	1	8	[9]	15
	#8	9	15	5	12	6	3	1	[8]
Newcomers	#9	1	5	12	9	6	15	8	3
	#10	6	1	12	9	5	15	3	8
	#11	15	5	9	1	12	6	8	3
	#12	8	15	3	5	6	9	12	1

allows for more accurate comparability between the two. Second, this variant allows us to compare the participation rates induced by the two mechanisms. Third, in real life some existing tenants might not return their applications, which has been treated as signing up for the same room, for example, at the University of Michigan.[7]

II. Experimental Design

We designed our experiment to compare outcomes of RSD and TTC mechanism, with particular attention to the questions of efficiency comparison and the participation decision of existing tenants. The environment was designed to capture the key aspects of the house allocation problem and to simulate the complexity inherent in potential applications. We implemented three different treatments: the original treatment and two other treatments for robustness tests. These will be explained in turn.

In the **original treatment,** there are 12 participants per session. Participants #1–#8 are existing tenants, each of whom currently lives in a house. Participants #9–#12 are newcomers, each of whom does not have a house yet. There are 12 houses of eight different types to allocate. Each house can only be allocated to one participant. Participants #1–#8 currently live in houses of types A–H. There are four additional vacant houses, one each of types A, B, C, and D. Therefore, there are two houses of the types A, B, C, D each and one house of the types E, F, G, H each. We choose to have eight existing tenants out of 12 participants based on two considerations: (1) In field applications involving dormitory assignment or office assignment, the majority of the agents are existing tenants. (2) Since the existing tenants' behavior is crucial for the efficiency of the mechanisms, we need a fairly large number of observations.

Table 1 presents the monetary payoff (i.e., the induced preferences over houses; Vernon L. Smith, 1982) for each participant as a result of the type of house she holds at the end of the experiment. A square bracket, [ ], indicates that the participant currently lives in a house of the specified type. For example, participant #1 lives in a house of type A. She will get $6 if she gets a house of type A at the end of the experiment, $3 if she gets a house of type B, etc. These payoff parameters are chosen with the following considerations:

1. There are nine Pareto-efficient house allocations for the chosen problem.[8] In general the aggregate payoff can differ at different Pareto-efficient allocations. We chose the payoff parameters such that the aggregate payoff is 171 at each of these efficient allocations. This conveniently gives us a unique reference point for full efficiency.

[7] See Michigan *Reapplication Lease Renewal Program* (University Housing, 1999, p. 2).

[8] In all Pareto-efficient allocations

(i) 1 gets E, 2 gets G, 3 gets D, 6 gets A, 7 gets H, 8 gets B, 11 gets A, 12 gets B,
(ii) 4, 5, 9, and 10 share two Cs, D, and F,
(iii) F is assigned to one of 5, 9, and 10.

2. To make the existing tenant's problem interesting, existing tenants' houses range from their second to the seventh choice. If the existing tenant's house is her first or eighth (i.e., last) choice, her decision to enter the lottery becomes trivial.
3. The payoffs between different outcomes are sufficiently dispersed so that there is a monetarily salient difference ($14) between getting one's top choice and last choice.

In this environment we test the two mechanisms: RSD and TTC. In both mechanisms, existing tenants are explicitly given an option to keep their houses and thus not enter the lottery.

Under RSD, truthful preference revelation is a dominant strategy for the newcomers. In case an existing tenant chooses to enter the lottery, truthful preference revelation is also a dominant strategy for her. In this case the outcome depends on the random ordering and she may end up with a worse house than she holds. Thus we expect to see some opting out.

Under TTC, truthful preference revelation is a dominant strategy for everyone. In addition individual rationality of the top trading cycles mechanism implies that opting out is a dominated strategy for existing tenants. Thus we expect to see existing tenants always participating.

Both mechanisms were implemented as one-shot games of incomplete information. Each subject knew her own payoff table, but not the other participants' payoff tables. They did know that "different participants might have different payoff tables." This information condition is a good approximation of reality. To check robustness, we also conducted an experiment to compare the performance of RSD and TTC when agents have complete information about all payoff tables. Results from that experiment can be found in Chen and Sönmez (1999). All qualitative results from the current experiment hold under complete information as well.

We use one-shot games to evaluate the mechanisms since in real-world applications the mechanisms will likely be used in a one-shot setting. While it is true that in college many students go through the house allocation process more than once, each time they face a different population, different sets of dormitory rooms, and different preferences. Hence, from a theoretical perspective, each house allocation process is a one-shot game. Without any practice rounds or opportunities to learn over time, one-shot implementation presents the most realistic and the toughest test for the mechanisms.[9]

To test robustness of the results with respect to changes in size and environments, we conducted two additional treatments. In the **large treatment** we replicate the economy in the original treatment by a factor of five, while keeping other features of the original treatment the same. Therefore, in the large treatment there are 60 participants per session. Participants #1–#40 are existing tenants, while participants #41–#60 are newcomers. There are 60 houses of eight different types to allocate. Each of participants #1–#5 currently lives in a house of type A; each of participants #6–#10 currently lives in a house of type B; ... ; and each of participants #36–#40 currently lives in a house of type H. There are 20 additional vacant houses, five each of types A, B, C, and D. Therefore, there are ten houses of the types A, B, C, D each, and five houses of the types E, F, G, H each. The instructions for the large treatment are identical to those for the original treatment except for the following two parts: (1) For both RSD and TTC the number of each type of participants and who lives where are changed accordingly. (2) For TTC, since now there are five of each type, in describing the house allocation method we added one sentence to describe which one of them moves to the head of the assignment queue once their house type is requested, "Among these existing tenants, the one who is closest to the top of the queue is given priority and moved to the top of the priority queue, directly in front of the requester." To assess the robustness of the results with respect to different parameter conditions, we conducted a third treatment, the **random treatment,** where we randomly generated a payoff matrix[10] subject to the following constraints: (1) For each agent the payoff for each type of housing is randomly drawn from {1, 2, ... , 15} without replacement; (2) Existing tenants' houses range from their second to the

[9] In experimental economics, one-shot implementations have been studied in various contexts (Reinhard Selten and Axel O. Ockenfels, 1998; Klaus Abbink et al., 2000).

[10] See Roth et al. (1998) for an example of using randomly generated environments for experiments.

TABLE 2—PAYOFF TABLE FOR ALL AGENTS—RANDOMLY GENERATED ENVIRONMENT

Type of houses:		A	B	C	D	E	F	G	H
Existing tenants	#1	[13]	9	8	14	10	15	4	1
	#2	11	[6]	1	13	4	3	12	10
	#3	15	9	[13]	6	10	1	14	2
	#4	12	9	11	[5]	10	7	2	15
	#5	2	14	5	15	[6]	3	10	12
	#6	4	13	15	12	6	[8]	2	10
	#7	15	5	4	1	12	11	[10]	9
	#8	11	5	2	13	7	3	15	[12]
Newcomers	#9	4	15	1	12	11	10	9	5
	#10	14	6	15	5	11	3	1	2
	#11	5	8	6	11	1	7	13	14
	#12	5	14	8	11	7	10	15	3

TABLE 3—FEATURES OF EXPERIMENTAL SESSIONS

Sessions	Dates	Mechanisms	Environments	Subjects per session	Total # of subjects
R_o1–R_o5	09/99	RSD	Original	12	60
T_o1–T_o5	09/99	TTC	Original	12	60
R_r1–R_r5	09/00	RSD	Random	12	60
T_r1–T_r5	09/00	TTC	Random	12	60
R_l1, R_l2	09/00	RSD	Original	60	120
T_l1, T_l2	09/00	TTC	Original	60	120

seventh choice, for the same reason as stated in choosing the original environment. Compared with the original treatment, the random treatment only differs in the payoff matrix, which is shown in Table 2.

For the original treatment five independent sessions for each mechanism (R_o1–R_o5, T_o1–T_o5) were conducted in September 1999 at the University of Michigan.[11] In September 2000 we conducted two independent sessions for each mechanism under the large treatment (R_l1, R_l2, T_l1, and T_l2), and five independent sessions for each mechanism (R_o1–R_o5, T_o1–T_o5) under the random treatment.

Table 3 summarizes features of experimental sessions, including dates experiments were conducted, environments, subjects per session, and total number of subjects under each treatment. All sessions were conducted by hand. Our subjects were undergraduate students at the University of Michigan. No subject participated in more than one session. This gives us a total of 24 independent sessions and 480 subjects. Each session consisted of one round only. The sessions lasted between 40–45 minutes, with the first 20–25 minutes being used for instructions. The conversion rate was $1 for all sessions. Each subject also received a participation fee of $3 in addition to their earnings from the experiment. The average earning (including participation fee) was $14.70.

Each subject randomly drew an envelope with an ID number and instructions inside, then was seated in a chair in a classroom. The experimenter read the instructions aloud. Subjects asked questions, which were answered in public. Subjects were then given ten minutes to read the instructions again at their own pace and to make their decisions. At the end of ten minutes the experimenter collected the decisions and asked volunteers to come to the front to draw

[11] There were also two incomplete sessions for RSD, in each of which one subject left in the middle of the experiment and the experimenter allocated all houses as if there were 12 houses and 11 participants. Since that might have changed the perception of the problem for others, we exclude these two sessions in the analysis. Results including these two sessions are not qualitatively different (they are in fact stronger) and are available from the authors upon request.

TABLE 4—EFFICIENCY—ORIGINAL ENVIRONMENT
(Standard Errors in Parentheses)

Mechanisms	Sessions	Observed efficiency	Expected efficiency with 1 million lotteries	Recombinant estimation of mean efficiency
RSD (original)	R_o1	0.673	0.693 (0.031)	$\hat{\mu}_{rsd}$ = 0.754 (0.020)
	R_o2	0.737	0.741 (0.023)	σ^2 = 0.00358
	R_o3	0.836	0.849 (0.027)	φ = 0.000203
	R_o4	0.661	0.698 (0.027)	
	R_o5	0.750	0.802 (0.033)	
RSD (large)	R_l1	0.743	0.742 (0.007)	$\hat{\mu}_{rsd}$ = 0.742 (0.001)
	R_l2	0.737	0.746 (0.013)	σ^2 = 0.000331
				φ = 0.0000000542
TTC (original)	T_o1	0.924	0.934 (0.061)	$\hat{\mu}_{ttc}$ = 0.889 (0.020)
	T_o2	0.743	0.802 (0.044)	σ^2 = 0.00332
	T_o3	0.901	0.871 (0.050)	φ = 0.000157
	T_o4	0.930	0.911 (0.021)	
	T_o5	0.877	0.890 (0.025)	
TTC (large)	T_l1	0.913	0.903 (0.010)	$\hat{\mu}_{ttc}$ = 0.875 (0.006)
	T_l2	0.837	0.830 (0.012)	σ^2 = 0.000692
				φ = 0.00000107

ping-pong balls out of an urn, which generated the initial priority queue. For the original and random treatments, the experimenter allocated all houses in public according to the specific house allocation mechanism, and paid the subjects at the end of the experiment. For the large treatment, because of the complexity involved in computing the results by hand, the experimental sessions ended after the initial priority queue was generated. Then the experimenter put the subjects' decisions and queues into a computer to generate the allocations, announced the allocations and paid the subjects the next day.

Experimental instructions and data are available from the authors upon request. Within each treatment the instructions for the two mechanisms are identical except for the "House Allocation Method" section. Across treatments, minimal changes in the instructions are made only in the few places as explained above.

III. Results

Three questions are important in evaluating the mechanisms. The first is the efficiency of the mechanisms. The second is whether individuals play their dominant strategies. The third is the robustness of the experimental results with respect to changes in the environment and the size of the economy. In subsection A, we report and analyze the experimental results. In subsection B, we report simulation results.

A. *Experimental Results*

To evaluate the aggregate performance of the mechanisms, we compare the efficiency generated by each mechanism. To extract the maximum information from the data, we look at three efficiency measures—observed efficiency, expected efficiency, and the recombinant estimation of mean efficiency.

Table 4 reports the efficiency estimates under all three measures for the original and large treatments. Table 5 reports the efficiency estimates under all three measures for the random treatment. *Observed efficiency* is calculated by taking the ratio of the sum of actual earnings of all subjects in a session and the Pareto-optimal earnings of the group. The unique Pareto-optimal group earning is 171 for the original treatment, and 855 for the large treatment. Since there are many Pareto-optimal allocations with different group payoffs under the random treatment, we use the Monte Carlo method to calculate the mean of efficient payoffs. That is, we randomly generate one million queues and let each participant pick her best choice among the

TABLE 5—EFFICIENCY—RANDOMLY GENERATED ENVIRONMENT
(Standard Errors in Parentheses)

Mechanisms	Sessions	Observed efficiency	Expected efficiency with 1 million queues	Recombinant estimation of mean efficiency
RSD (random)	R_r1	0.865	0.875 (0.023)	$\hat{\mu}_{rsd} = 0.877\ (0.001)$
	R_r2	0.858	0.876 (0.023)	$\sigma^2 = 0.00057$
	R_r3	0.901	0.884 (0.023)	$\varphi = 0.000000547$
	R_r4	0.834	0.876 (0.023)	
	R_r5	0.913	0.873 (0.026)	
TTC (random)	T_r1	0.919	0.934 (0.018)	$\hat{\mu}_{ttc} = 0.931\ (0.018)$
	T_r2	0.865	0.877 (0.028)	$\sigma^2 = 0.00252$
	T_r3	0.925	0.947 (0.031)	$\varphi = 0.000132$
	T_r4	1.041	1.000 (0.027)	
	T_r5	0.938	0.925 (0.018)	

available houses down the queue and calculate the average payoff over one million assignments, which is 164.24.

RESULT 1 (Observed Efficiency): *The observed efficiency of TTC is significantly higher than that of RSD.*

SUPPORT:

Column 3 in Tables 4 and 5 present the observed efficiency of the two mechanisms. Permutation tests[12] show that the observed efficiency of TTC is significantly higher than that of RSD: $p = 0.0079$ (one-tailed) for the original treatment, $p < 0.0001$ (one-tailed) for the large treatment, $p = 0.0198$ (one-tailed) for the random treatment.

Given the participation decisions and the participants' stated preference orderings, the observed efficiency is dependent upon the realization of the random selection of the queue in which the mechanism assigns those participants. In an experimental session we observe only one out of $n!$ possible queues, where n is the number of participants entering the lottery. To get an idea of the *expected efficiency* in each session, we calculate the average efficiency over one million randomly generated lotteries within each session.

RESULT 2 (Expected Efficiency): *The expected efficiency of TTC is significantly higher than that of RSD.*

SUPPORT:

Column 4 in Tables 4 and 5 present the expected efficiency of the two mechanisms. Permutation tests show that the expected efficiency of TTC is significantly higher than that of RSD: $p = 0.004$ (one-tailed) for the original treatment, $p < 0.0001$ (one-tailed) for the large treatment, and $p = 0.004$ (one-tailed) for the random treatment.

Since both mechanisms are implemented as true one-shot games with a total of 480 independent observations, we can use an improved statistical estimator, the recombinant estimator (David Lucking-Reiley and Charles Mullin, 1999), to compare the *mean efficiency*. With recombinant estimation, one recombines the strategies of different players to compute what the outcomes would have been if players' grouping had been different. In RSD or TTC we have a total of $(5)^{12}$ different groups under the original and random treatment, and $(10!/(5!5!))^{12}$ different groups under the large treatment. In the actual estimation we randomly generated two million groups (and one queue

[12] The permutation test, also known as the Fisher randomization test, is a nonparametric version of a difference of two means t-test (see, e.g., Sidney Siegel and N. John Castellan, Jr., 1988, pp. 95–100). By pooling the ten independent observations, the p-value is obtained as the exact probability of observing a separation between the two treatments as the one observed when the pooled observations are randomly divided into two equal-sized groups. This test uses all of the information in the sample, thus has power efficiency of 100 percent.

TABLE 6—EXISTING TENANTS' PARTICIPATION RATES

Mechanisms	Original sessions	Participation rates	Random sessions	Participation rates	Large sessions	Participation rates
RSD	R_o1	3/8	R_r1	4/8	R_l1	16/40
	R_o2	4/8	R_r2	4/8	R_l2	19/40
	R_o3	5/8	R_r3	4/8	—	—
	R_o4	3/8	R_r4	4/8	—	—
	R_o5	5/8	R_r5	4/8	—	—
	Overall	0.500	Overall	0.500	Overall	0.438
TTC	T_o1	7/8	T_r1	6/8	T_l1	34/40
	T_o2	5/8	T_r2	5/8	T_l2	30/40
	T_o3	6/8	T_r3	6/8	—	—
	T_o4	7/8	T_r4	8/8	—	—
	T_o5	6/8	T_r5	6/8	—	—
	Overall	0.775	Overall	0.775	Overall	0.800

for each group) under each mechanism to estimate the mean ($\hat{\mu}$), variance (σ^2), and covariance (φ) of the data.

RESULT 3 (Mean Efficiency): *The mean efficiency of TTC is significantly higher than that of RSD. The recombinant estimation of mean efficiency is* 88.9 *percent (original),* 87.5 *percent (large), and* 93.1 *percent (random) for TTC, and* 75.4 *percent (original),* 74.2 *percent (large), and* 87.7 *percent (random) for RSD.*

SUPPORT:

The last column of Tables 4 and 5 reports the recombinant estimation of the mean efficiency of the two mechanisms. A t-test of H_0: $\hat{\mu}_{ttc} = \hat{\mu}_{rsd}$ against H_1: $\hat{\mu}_{ttc} > \hat{\mu}_{rsd}$ yields $z = 4.60$ ($p < 0.001$) under the original treatment, $z = 2.28$ ($p = 0.0113$) under the large treatment, and $z = 3.04$ ($p = 0.0012$) under the random treatment.

Therefore, by all three measures of efficiency TTC outperforms RSD. There are two sources for the loss of efficiency: nonparticipation of existing tenants; and manipulation of preferences. Next we examine the participation decisions and types of preference orderings observed in the experiment.

RESULT 4 (Participation): *Existing tenants under TTC are significantly more likely to participate than those under RSD. The existing tenants' overall participation rate is* 78.8 *percent under TTC, but only* 46.9 *percent under RSD.*

SUPPORT:

Table 6 presents existing tenants' participation rates for each session under all three treatments. T-tests of proportions show that the participation rate of existing tenants under TTC is significantly higher than that of RSD: $z = 2.558$ ($p = 0.0052$, one-tailed) under the original treatment, $z = 4.720$ ($p < 0.001$, one-tailed) under the large treatment, and $z = 2.558$ ($p = 0.0052$, one-tailed) under the random treatment.

Therefore, existing tenants' participation rate is significantly higher under TTC than under RSD. We next examine the size and environment effects on participation rates.

RESULT 5 (Size and Environment Effects on Participation): *Under each mechanism, the differences in participation rates between the original treatment and the large treatment, and between the original and the random treatment are not statistically significant.*

SUPPORT:

Table 6 presents existing tenants' participation rates for each session under all three treatments. T-tests of proportions show that under RSD, participation rates under the original treatment are not significantly different from those under the large treatment: $z = 0.648$ ($p = 0.7422$, one-tailed); under RSD, participation rates under the original treatment are not significantly different from those under the random treatment: $z = 0.000$ ($p = 0.5000$, one-tailed). Similar results hold under TTC: $z = 0.318$ ($p = 0.6255$,

TABLE 7—OPTING OUT AND OWN-HOUSE PAYOFFS, v_o

	RSD_o		RSD_l			RSD_r		RSD overall	
Payoff	# of ET	# Out	# of ET	# Out	Payoff	# of ET	# Out	Range	Percentage out
12	10	10	20	20	13	10	10		
9	10	6	20	16	12	5	5	≥10	100
8	5	2	10	4	10	5	5	8, 9	56
6	5	2	10	2	8	5	0	6, 7	16
5	5	0	10	3	6	10	0	4, 5	15
3	5	0	10	0	5	5	0	≤3	0

	TTC_o		TTC_l			TTC_r		TTC overall	
Payoff	# of ET	# Out	# of ET	# Out	Payoff	# of ET	# Out	Range	Percentage out
12	10	8	20	9	13	10	2		
9	10	1	20	5	12	5	3	≥12	49
8	5	0	10	1	10	5	1	10, 11	20
6	5	0	10	1	8	5	1	8, 9	16
5	5	0	10	0	6	10	2	6, 7	12
3	5	0	10	0	5	5	0	≤5	0

one-tailed) between the original and large treatments, and $z = 0.000$ ($p = 0.5000$, one-tailed) between the original and random treatments.

Result 5 is important because it demonstrates that existing tenants' participation rates induced by each mechanism are robust with respect to variations in size and environment.

Although existing tenants' participation rate is significantly higher under TTC than under RSD, it is still 21 percent below the theoretically predicted 100-percent participation rate under TTC. We now explore the possibility that various factors might contribute to individual's tendency to participate under each mechanism. We first examine the effects of payoffs from existing tenants' own houses (v_o) on participation decisions.

Table 7 presents the number and proportion of existing tenants (ET) who opted out and the payoffs they received from their own houses. Under both mechanisms, as the payoff from own house increases, the proportion of opting out also increases. There are two possible explanations for why some existing tenants opted out under TTC. First, since TTC is relatively complex, it requires some effort, or decision cost, to understand the mechanism. Some subjects, content with what they could earn by opting out, decided not to bother to try to understand the instructions and thus avoided the decision cost. This explains why we observe more opting out with higher payoffs. Second, we cannot rule out that some subjects incurred the decision cost but still did not understand how the TTC mechanism works. This might explain why we observe some opting out in low-payoff ranges. In RSD, however, since the mechanism itself is not individually rational, many more existing tenants with a wider range of payoffs opted out.

An alternative model says that existing tenants' participation decision is affected by the difference of the average values of all houses ($\bar{v} = (2(v_a + v_b + v_c + v_d) + (v_e + v_f + v_g + v_h))/12$), and own-house value ($v_o$), $d = \bar{v} - v_o$. This is a "rule-of-thumb" type of model. Payoff from opting out is v_o; however, expected payoffs from opting in is much more complicated to compute, since in RSD there are typically multiple equilibria, while in TTC a participant might also consider the average payoff if he did not understand the mechanism. Table 8 tabulates the number and proportion of existing tenants who opted out under each mechanism in various ranges of d.

In order to test the predictive capacity of the two simple models, we do a logit analysis. In the logit analysis the dependent variable is a discrete choice variable, IN, which equals one if

TABLE 8—OPTING OUT AND PAYOFF DIFFERENCES, $d = \bar{v} - v_o$

		RSD			TTC		
$d = \bar{v} - v_o$	v_o	# of ET	# out	Percentage out	# of ET	# out	Percentage out
≤ -3	13; 12	45	45	100	45	22	49
$(-3, -2]$	10	5	5	100	5	1	20
$(-2, 0]$	9	30	22	73	30	6	20
$(0, 1]$	(o)8	15	6	40	15	1	7
$(1, 3]$	(o)5; (r)8; 6	45	7	16	45	4	9
> 3	(r)5, 3	20	0	0	20	0	0

TABLE 9—LOGIT MODELS OF PARTICIPATION DECISIONS

Independent variables	RSD_{v_o}	RSD_d	TTC_{v_o}	TTC_d
Constant	7.329	0.219	5.139	1.992
Standard error	(1.163)***	(0.262)	(0.914)***	(0.320)***
Own-house value (v_o)	−0.920		−0.412	
Standard error	(0.141)***		(0.088)***	
Difference (d)		1.000		0.490
Standard error		(0.139)***		(0.109)***
Log-likelihood	−53.691	−49.731	−67.637	−67.410
Percent correctly predicted	81.875	86.875	75.000	78.750

*** Significant at the 1-percent level.

an existing tenant opted in, and zero otherwise. We consider two independent variables, the value of own house, v_o, and the difference between the average value of all houses and the value of own house, d. Therefore, the model is

$$P[\text{IN} = 1] = \Phi(\boldsymbol{\beta}'\mathbf{x}).$$

For RSD_{v_o} and TTC_{v_o} the only independent variable is the value of own house, v_o. For RSD_d and TTC_d the only independent variable is the difference between the average value of all houses and value of own house, d. Coefficients, standard errors, log-likelihood and the percentages correctly predicted for each model are given in Table 9.

A consistent pattern in both RSD_d and TTC_d is the positive and significant impact of d on the participation decisions. From the last two lines, log-likelihood and percentage correctly predicted, RSD_d fits the data much better than RSD_{v_o}, while TTC_d fits the data slightly better than TTC_{v_o}. Therefore, we will base the simulation analysis in the next subsection on Table 8.

Once the existing tenant decides to participate, then truthful preference revelation is a dominant strategy for both TTC and RSD. However, as many experiments have shown, subjects do not always play dominant strategies.

RESULT 6 (Truthful Preference Revelation): *The overall proportion of truthful preference revelation is* 70.9 *percent under TTC, and* 73.5 *percent under RSD. The differences in the proportions of truthful preference revelation under TTC and RSD are not statistically significant.*

SUPPORT:

Table 10 shows the proportion of truthful preference revelation for each session under both mechanisms. T-tests of proportions show that the proportion of truthful preference revelation under RSD is not significantly different from that of TTC: $z = 0.141$ ($p = 0.5557$) under the original treatment, $z = 0.478$ ($p = 0.6808$) under the large treatment, and $z = 0.256$ ($p = 0.5987$) under the random treatment.

A little over two-thirds of the subjects revealed their preferences truthfully, i.e., ranked

TABLE 10—PROPORTION OF TRUTHFUL PREFERENCE REVELATION

Mechanisms	Original sessions	Proportion of truth	Random sessions	Proportion of truth	Large sessions	Proportion of truth
RSD	R_o1	5/7	R_r1	6/8	R_l1	27/36
	R_o2	5/8	R_r2	7/8	R_l2	26/39
	R_o3	8/9	R_r3	7/8	—	—
	R_o4	4/7	R_r4	7/8	—	—
	R_o5	6/9	R_r5	6/8	—	—
	Overall	0.700	Overall	0.825	Overall	0.707
TTC	T_o1	7/11	T_r1	9/10	T_l1	35/54
	T_o2	5/9	T_r2	8/9	T_l2	35/50
	T_o3	9/10	T_r3	6/10	—	—
	T_o4	9/11	T_r4	11/12	—	—
	T_o5	5/10	T_r5	7/10	—	—
	Overall	0.682	Overall	0.801	Overall	0.673

TABLE 11—CLASSIFICATION OF MISREPRESENTATION
(Proportion in Parentheses)

	RSD_o	RSD_r	RSD_l	RSD total	TTC_o	TTC_r	TTC_l	TTC total
Vacant house	8	4	12	24 (0.155)	13	8	21	42 (0.204)
Switch-top-two	1	0	1	2 (0.013)	2	0	3	5 (0.024)
Switch-lower-two	0	1	0	1 (0.006)	0	2	0	2 (0.010)
Duplicate	2	0	0	2 (0.013)	0	0	0	0 (0.000)
Random	1	2	9	12 (0.077)	1	0	10	11 (0.053)
Total misrepresented	12	7	22	41 (0.265)	16	10	34	60 (0.291)
Total truth	28	33	53	114 (0.735)	35	41	70	146 (0.709)

the houses in terms of decreasing payoffs. Among the subjects who misrepresented their preferences, we can classify the misrepresentation in five categories. Table 11 presents the number of cases and proportion of misrepresentation under each of the five categories.

Nearly two-thirds of the misrepresentation is due to the *vacant house effect*—since there are four (or 20) vacant houses, one (or five) each of types A, B, C, and D, some subjects switched their choices in favor of types A, B, C, or D over the rest of the house types. In seven cases (*switch-top-two*), subjects switched their top two choices which did not belong to the case of the vacant house effect. This might be due to a perception that people might have similar preferences and therefore their top choice might be more competitive. In three cases (*switch-lower-two*) subjects switched their lower two choices which did not belong to the case of the vacant house effect. In two cases (*duplicate ranking*), the rankings were monotone decreasing in payoffs; however, the subjects ranked A (or B or C or D) twice, thus the rankings were incomplete.[13] In nearly a quarter of the misrepresentation cases (*random*) the ranking seemed totally random. It might be due to the subjects' confusion. Comparing RSD_o and RSD_l, TTC_o and TTC_l, a noticeable size effect on the pattern of misrepresentation is that the proportions of the Random category increased by a factor of five.

RESULT 7 (Size and Environment Effects on Truth-telling): *Under both mechanisms, the difference in truth-telling rates between the original treatment and the large treatment is not statistically significant, while the proportion of truth-telling under the random treatment is*

[13] In Monte Carlo simulations, we complete the ranking for these two cases in terms of decreasing payoffs, since the original ranking was monotone decreasing. Results 1 to 3 do not change if we complete the ranking by randomly generating a list.

TABLE 12—VACANT HOUSE MANIPULATION

1st choice	2nd choice	Total # of agents	Total # in queue	# of vacant house manipulation	Percent of vacant house manipulation
full	full	70	63	11	17
full	vacant	160	130	33	50
vacant	full	150	101	16	24
vacant	vacant	100	67	6	9
	Total	480	361	66	100

weakly higher than that under the original treatment.

SUPPORT:

Table 10 shows the proportion of truthful preference revelation for each session under both mechanisms. T-tests of proportions show that the difference in truth-telling rates between the original treatment and the large treatment is not statistically significant: $z = 0.0746$ ($p = 0.5279$) under RSD; and $z = 0.1652$ ($p = 0.5675$) under TTC. However, the proportion of truth-telling under the random environment is weakly higher than that under the original environment: $z = 1.3136$ ($p = 0.0951$) under RSD; and $z = 1.3632$ ($p = 0.0869$) under TTC.

Result 7 says that under both mechanisms the size effect on truth-telling is not statistically significant. However, we do observe a weakly significant environment effect. If we compare columns RSD_o with RSD_r, and TTC_o with TTC_r in Table 11, there is a notable decrease of the number of misrepresentations due to the vacant house effect, which accounts for nearly two-thirds of the misrepresentations. This might be due to the properties of the payoff matrices: in the original environment (Table 1), 50 percent of the top two choices are vacant (i.e., of types A, B, C, or D), while in the random environment (Table 2), 63 percent of the top two choices are vacant. If a participant's top choices are already vacant, there is no need for the vacant house manipulation.

Table 12 tabulates the proportion of agents doing vacant house manipulations as a function of different types of first and second choices. Over 90 percent of the agents with vacant house manipulations have at least one of their top choices full. Only 9 percent of these agents have both of their top choices vacant.

Having analyzed the size and environment effects on participation and truth-telling, we now examine the size and environment effects on efficiency.

RESULT 8 (Size Effect on Efficiency): *Under each mechanism, the difference in efficiency under the original treatment and the large treatment is not statistically significant.*

SUPPORT:

Table 4 presents all three efficiency measures under the two treatments. Permutation tests show that under RSD observed and expected efficiencies under the original treatment are not significantly different from those under the large treatment: $p = 0.4286$ (one-tailed) for observed efficiency and $p = 0.4762$ (one-tailed) for expected efficiency. A t-test of H_0: $\hat{\mu}_{rsd_o} = \hat{\mu}_{rsd_l}$ against H_1: $\hat{\mu}_{rsd_o} > \hat{\mu}_{rsd_l}$ yields $z = 0.526$ ($p = 0.7019$). Permutation tests also show that under TTC observed and expected efficiencies under the original treatment are not significantly different from those under the large treatment: $p = 0.4286$ (one-tailed) for observed efficiency and $p = 0.3810$ (one-tailed) for expected efficiency. A t-test of H_0: $\hat{\mu}_{ttc_o} = \hat{\mu}_{ttc_l}$ against H_1: $\hat{\mu}_{ttc_o} > \hat{\mu}_{ttc_l}$ yields $z = 0.706$ ($p = 0.7611$).

Result 8 shows that a fivefold increase in the size of the economy does not lead to significantly different performance of either mechanism. However, one needs to be cautious to extrapolate from Result 8 to an increase of size beyond 60 subjects per session.

RESULT 9 (Environment Effect on Efficiency): *Under RSD, efficiency under the random treatment is significantly higher than*

TABLE 13—SIMULATION RESULTS: MEAN EFFICIENCY
(Standard Errors in Parentheses)

	n = 12		n = 60		n = 300		n = 900		n = 1,500	
Mechanisms	RSD	TTC	RSD	TTC	RSD	TTC	RSD	TTC	RSD	TTC
Calibration	0.878	0.941	0.822	0.915	0.806	0.909	0.807	0.911	0.807	0.911
With observed behavior	(0.039)	(0.020)	(0.042)	(0.020)	(0.044)	(0.020)	(0.041)	(0.016)	(0.041)	(0.016)
25 percent random	0.850	0.880	0.786	0.852	0.758	0.843	0.765	0.847	0.761	0.843
Preference ordering	(0.028)	(0.021)	(0.031)	(0.014)	(0.036)	(0.016)	(0.030)	(0.014)	(0.030)	(0.013)
50 percent random	0.796	0.802	0.727	0.766	0.702	0.757	0.704	0.757	0.702	0.755
Preference ordering	(0.030)	(0.027)	(0.023)	(0.012)	(0.024)	(0.011)	(0.018)	(0.010)	(0.020)	(0.009)
70 percent random	0.754	0.741	0.680	0.700	0.658	0.689	0.657	0.688	0.656	0.686
Preference ordering	(0.035)	(0.030)	(0.020)	(0.012)	(0.014)	(0.007)	(0.011)	(0.006)	(0.012)	(0.005)
Zero intelligence	0.638	0.638	0.608	0.607	0.590	0.590	0.587	0.587	0.584	0.584
	(0.042)	(0.046)	(0.017)	(0.018)	(0.009)	(0.009)	(0.005)	(0.005)	(0.002)	(0.002)
Education	0.906	1.000	0.849	1.000	0.833	1.000	0.829	1.000	0.837	1.000
	(0.036)	(0.013)	(0.041)	(0.004)	(0.051)	(0.001)	(0.040)	(0.001)	(0.047)	(0.001)

that under the original environment. Under TTC, efficiency under the random treatment is weakly higher than that under the original environment.

SUPPORT:
Tables 4 and 5 present all three efficiency measures under the original and random treatments respectively. Permutation tests show that under RSD observed and expected efficiencies under the random treatment are significantly higher than those under the original treatment: $p = 0.0079$ (one-tailed) for observed efficiency and $p = 0.0040$ (one-tailed) for expected efficiency. A t-test of $H_0{:}\hat{\mu}_{rsd_r} = \hat{\mu}_{rsd_o}$ against $H_1{:}\hat{\mu}_{rsd_r} > \hat{\mu}_{rsd_o}$ yields $z = 5.57$ ($p < 0.001$). Permutation tests show that under TTC observed and expected efficiencies under the random treatment are weakly higher than those under the original treatment: $p = 0.1151$ (one-tailed) for observed efficiency and $p = 0.0556$ (one-tailed) for expected efficiency. A t-test of $H_0{:}\hat{\mu}_{ttc_o} = \hat{\mu}_{ttc_r}$ against $H_1{:}\hat{\mu}_{ttc_r} > \hat{\mu}_{ttc_o}$ yields $z = 1.60$ ($p = 0.0548$).

Result 9 shows that there is an environment effect on efficiency. Over all three treatments the efficiency under TTC is significantly higher than that under RSD, although the magnitude of the difference in efficiency might vary when the environments are different. To explore the performance of the two mechanisms over a large number of different environments, different scale of the market, and different strategy distributions, we resort to simulation analysis.

B. *Simulation Results*

In this subsection we report simulation results with varying sizes of the economy, environments, as well as the strategies inputted into the simulation. In each simulation treatment we randomly generate 50 payoff matrices to further explore the robustness of the results with respect to changes in environments. We also significantly increase the scale of the market. We simulate the nature of the participation decisions using the logit participation rule based on the difference model that was estimated from the actual data and reported in Table 8.

Table 13 reports the mean efficiency over 50 randomly generated environments under each mechanism for each simulation treatment. In the first set of simulations, we use the same percentages on participation (Table 8) and the same distribution over types of preference orderings (Table 11) as observed in the actual experiment. We then vary the scale of the market (n = 12, 60, 300, 900, 1,500) and use 50 randomly generated payoff matrices for each n to check whether the same efficiency ranking of the two mechanisms would be preserved. In each Monte Carlo simulation, we use 1,000 randomly generated queues[14] to calculate the Pareto-optimal payoffs as well as the agents' aggregate payoffs from each environment and for each mechanism.

RESULT 10 (Calibration Based on Observed Behavior): *As the scale of market increases, the mean efficiency of TTC over* 50 *randomly generated environments remains significantly higher than that of RSD.*

SUPPORT:
Rows 3 and 4 of Table 13 report the mean efficiency and standard error over 50 randomly generated environments for each mechanism and each n. T-tests of each pair yield $p < 0.001$.

Result 10 shows that the same efficiency ranking of the two mechanisms is preserved as we significantly increase the scale of the market and calibrate based on the actual observed behavior. Indeed, it is preserved under each of the 250 random environments. In the next two sets of simulations, strategies inputted into the simulations go beyond those actually submitted by the participants of the experiments.

In the second set of simulations, we examine the performance of the two mechanisms as the proportion of random decisions increase. From row 5 to row 12, we increase the proportion of random preference orderings and decrease the proportion of truth-telling accordingly, while keeping the same participation decisions as those actually observed in the experiment. In rows 11 and 12 (zero-intelligence[15]), all decisions, preference ranking, and participation are completely random. This provides a benchmark for the performance of the two mechanisms.

RESULT 11 (Effects of Random Decision-making): *As the proportion of random decisions increases, the performance of both mechanisms decreases, and the gap between the performance of TTC and RSD decreases.*

SUPPORT:
From row 3 to row 11 of Table 13, for each n and each mechanism the efficiency decreases, while for each n the gap between TTC and RSD also decreases. The efficiency of the two mechanisms becomes indistinguishable with zero-intelligence participants (row 11).

In the third set of simulations, we explore the effects of education. There are two sources for the loss of efficiency: (1) nonparticipation of existing tenants; and (2) manipulation of preferences. In actual implementation of RSD the latter can be avoided by educating the agents while the former cannot be avoided by education. In contrast, in TTC all inefficiency can be avoided by education. Therefore, in the simulations we separate the two effects in RSD by using the actual participation data while assuming truthful preference revelation. In TTC we use truthful preference revelation and 100-percent participation rate.

RESULT 12 (Education): *The loss of efficiency in the actual experiment under RSD is mainly due to the nonparticipation behavior of existing tenants. Therefore, with education the efficiency of RSD improves by at most* 3.6 *percent while the efficiency of TTC reaches* 100 *percent.*

SUPPORT:
The last two rows in Table 13 report the efficiency under education. Comparing row 3 to row 13 we see only a small improvement in the efficiency of RSD, and a significant improvement under TTC.

Result 12 shows that in the actual implementation of the two mechanisms, if the planner (e.g., the housing office) can educate the participants effectively, we do not expect the efficiency of RSD to improve by a significant amount, while the efficiency of TTC can potentially reach 100 percent.

Overall, TTC is more efficient than RSD because it induces significantly higher participation rate of the existing tenants. These

[14] We compared the results with those using 10,000 queues and found no significant difference.

[15] See Dhananjay K. Gode and Shyam Sunder (1993) for using zero-intelligence traders in a market experiment.

results suggest that replacing RSD with TTC in practice can significantly improve allocation efficiency.

IV. Concluding Remarks

House allocation problems exemplify an interesting class of theoretical mechanism design problems, which also have important practical implications.

Currently many universities use some variant of the random serial dictatorship with squatting rights (RSD) mechanism, which is strategy-proof but not individually rational, and hence not Pareto efficient either. A simple mechanism, the top trading cycles (TTC) mechanism, was proposed by Abdulkadiroğlu and Sönmez (1999) as an alternative to RSD. The top trading cycles mechanism is not only strategy-proof, but also individually rational and Pareto efficient. In order to evaluate the actual performance of TTC relative to RSD when they are implemented among boundedly rational individuals, we designed an experiment to test these mechanisms in an environment which preserves all interesting features of the problem.

Our experimental evidence shows the TTC mechanism produces significantly more efficient allocations than the RSD mechanism. The mean efficiency of the TTC mechanism is 88 percent of full efficiency. In contrast, RSD only achieves 74 percent of full efficiency. The average efficiency gain is about 18 percent in our environment. The difference between the mechanisms is explained by the fact that significantly more existing tenants chose to enter the applicant pool under the TTC mechanism, as predicted by theory. Under each mechanism a little more than two-thirds of the subjects truthfully revealed their preferences. Our result suggests that replacing the random serial dictatorship with squatting rights mechanism with the top trading cycles mechanism in practice can significantly improve allocation efficiency.

REFERENCES

Abbink, Klaus; Irlenbusch, Bernd and Renner, Elke. "The Moonlighting Game: An Experimental Study on Reciprocity and Retribution." *Journal of Economic Behavior and Organization*, June 2000, *42*(2), pp. 265–77.

Abdulkadiroğlu, Atila and Sönmez, Tayfun. "Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems." *Econometrica*, May 1998, *66*(3), pp. 689–701.

______. "House Allocation with Existing Tenants." *Journal of Economic Theory*, October 1999, *88*(2), pp. 233–60.

Chen, Yan. "Incentive-Compatible Mechanisms for Pure Public Goods: A Survey of Experimental Research," in C. Plott and V. L. Smith, eds., *The handbook of experimental economics results*. Amsterdam: Elsevier Press, 2002 (forthcoming).

Chen, Yan and Sönmez, Tayfun. "An Experimental Study of House Allocation Mechanisms." Working paper, University of Michigan, 1999.

Demange, Gabrielle and Gale, David. "The Strategy Structure of Two-Sided Matching Markets." *Econometrica*, July 1985, *53*(4), pp. 873–88.

Gale, David and Shapley, Lloyd S. "College Admissions and the Stability of Marriage." *American Mathematical Monthly*, January 1962, *69*(1), pp. 9–15.

Gode, Dhananjay K. and Sunder, Shyam. "Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality." *Journal of Political Economy*, February 1993, *101*(1), pp. 119–37.

Harrison, Glenn W. and McCabe, Kevin A. "Stability and Preference Distortion in Resource Matching: An Experimental Study of the Marriage Problem," in R. Mark Isaac, ed., *Research in experimental economics*, Vol. 6. Greenwich, CT: JAI Press, 1996, pp. 53–129.

Haruvy, Ernan; Roth, Alvin E. and Ünver, M. Utku. "Dynamics of Law Clerk Matching: An Experimental and Computational Investigations of Proposals for Reform of the Market." Harvard NOM Research Paper No. 01-08, 2001.

Hylland, Aanund and Zeckhauser, Richard. "The Efficient Allocation of Individuals to Positions." *Journal of Political Economy*, April 1979, *87*(2), pp. 293–314.

Kagel, John H. "Auctions: A Survey of Experimental Research," in J. H. Kagel and A. E.

Roth, eds., *Handbook of experimental economics*. Princeton, NJ: Princeton University Press, 1995, pp. 501–85.

Kagel, John H. and Roth, Alvin E. "The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by a Natural Experiment." *Quarterly Journal of Economics*, February 2000, *115*(1), pp. 201–35.

Leonard, Herman B. "Elicitation of Honest Preferences for the Assignment of Individuals to Positions." *Journal of Political Economy*, June 1983, *91*(3), pp. 461–79.

Lucking-Reiley, David and Mullin, Charles. "Recombinant Estimation for Simultaneous-Move Games: An Application to Revenues and Allocations in Auctions." Working paper, Vanderbilt University, 1999.

Ma, Jinpeng. "Strategy-Proofness and the Strict Core in a Market with Indivisibilities." *International Journal of Game Theory*, April 1994, *23*(1), pp. 75–83.

Nalbantian, Haig R. and Schotter, Andrew. "Matching and Efficiency in the Baseball Free-Agent System: An Experimental Examination." *Journal of Labor Economics*, January 1995, *13*(1), pp. 1–31.

Olson, Mark and Porter, David. "An Experimental Examination into the Design of Decentralized Methods to Solve the Assignment Problem with and without Money." *Economic Theory*, January 1994, *4*(1), pp. 11–40.

Roth, Alvin E. "Incentive Compatibility in a Market with Indivisibilities." *Economics Letters*, February 1982, *9*(2), pp. 127–32.

______. "The Evolution of Labor Market for Medical Interns and Residents: A Case Study in Game Theory." *Journal of Political Economy*, December 1984, *92*(6), pp. 991–1016.

______. "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*, June 1991, *81*(3), pp. 441–64.

______. "Introduction to Experimental Economics," in J. H. Kagel and A. E. Roth, eds., *Handbook of experimental economics*. Princeton, NJ: Princeton University Press, 1995, pp. 3–109.

Roth, Alvin E.; Erev, Ido and Slonim, Robert. "Learning and Equilibrium as Useful Approximations: Accuracy of Prediction on Randomly Selected Constant Sum Games." Working paper, Harvard University, 1998.

Roth, Alvin E. and Postlewaite, Andrew. "Weak versus Strong Domination in a Market with Indivisible Goods." *Journal of Mathematical Economics*, August 1977, *4*(2), pp. 131–37.

Selten, Reinhard and Ockenfels, Axel O. "An Experimental Solidarity Game." *Journal of Economic Behavior and Organization*, March 1998, *34*(4), pp. 517–39.

Shapley, Lloyd S. and Scarf, Herbert E. "On Cores and Indivisibility." *Journal of Mathematical Economics*, March 1974, *1*(1), pp. 23–28.

Siegel, Sidney and Castellan, N. John, Jr. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1998.

Smith, Vernon L. "Microeconomic Systems as an Experimental Science." *American Economic Review*, December 1982, *72*(5), pp. 923–55.

Sönmez, Tayfun and Ünver, M. Utku. "House Allocation with Existing Tenants: An Equivalence." Working paper, Koç University, 2001.

University Housing. "Reapplication Lease Renewal Program 1999–2000." University of Michigan, 1999.

Ünver, M. Utku. "On the Survival of Some Two-Sided Matching Mechanisms: An Experimental and Computational Investigation of the Stability Hypothesis." Working paper, Koç University, 2001.

Zhou, Lin. "On a Conjecture by Gale about One-Sided Matching Problems." *Journal of Economic Theory*, October 1990, *52*(1), pp. 123–35.