

# House Allocation with Existing Tenants<sup>1</sup>

Atila Abdulkadiroğlu

*Department of Economics, University of Rochester, Rochester, New York 14627*

and

Tayfun Sönmez

*Department of Economics, University of Michigan, Ann Arbor, Michigan 48109-1220  
and Department of Economics, Koç University, 80860 Istanbul, Turkey*

Received October 14, 1998; revised April 29, 1999

In many real-life applications of house allocation problems, whenever an existing tenant wants to move, he needs to give up his current house before getting another one. This practice discourages existing tenants from such attempts and results in loss of potentially large gains from trade. Motivated by this observation, we propose a simple mechanism that is Pareto efficient, individually rational, and strategy-proof. Our approach is constructive and we provide two algorithms, each of which can be used to find the outcome of this mechanism. One additional merit of this mechanism is that it can accommodate any hierarchy of seniorities. *Journal of Economic Literature* Classification Numbers: C71, C78, D71, DF8. © 1999 Academic Press

## 1. INTRODUCTION

A class of resource allocation problems that is not only of theoretical interest, but also of practical importance is the class of *house allocation problems*: There is a set of houses (or some other indivisible items such as offices, tasks, etc.) which have to be allocated to a group of agents. Rents

<sup>1</sup> We are grateful to Haluk Ergin, Yves Sprumont, William Thomson, and an anonymous referee for their extensive comments. We would like to thank Julie Cullen, Glenn Ellison, Roger Gordon, John Knowles, Semih Koray, Bahar Leventoğlu, Benny Modovanu, Hervé Moulin, Tuvana Pastine, Phil Reny, Lones Smith, Ennio Stacchetti, Lars Stole, Asher Wolinsky, seminar participants at Bilkent, Carnegie-Mellon (GSIA), Michigan, Northwestern (MEDS), Rochester, SMU, 1998 North American Summer Meetings of the Econometric Society at Montreal, and the Forth International Meeting of the Society of Social Choice and Welfare at Vancouver for their helpful remarks. Sönmez gratefully acknowledges the research support of the National Science Foundation via Grants SBR-9709138 and SBR-9904214. All errors are our own responsibility.

are exogenously given and there is no medium of exchange such as money. While this assumption is restrictive, it is satisfied in many real-life applications, especially when housing is subsidized. Examples are on-campus housing and public housing. It is also a natural assumption if the indivisible good is used for job related purposes. Two such applications are office space and parking space allocated to university faculty. Usually there are several types of housing and the attractiveness of each type changes from person to person. Therefore the central planner, say the housing office, needs to find a “mechanism” to allocate houses to agents. A mechanism, namely the *random serial dictatorship*, and its variants are commonly used in many real-life applications of these problems. This mechanism randomly orders the agents and the first agent in the order is assigned his top choice, the next agent is assigned his top choice among the remaining houses, and so on. (In many applications the ordering is not entirely random and it depends on seniority as well.) In the original model where the sets of agents and houses are exogenously given, the random serial dictatorship has some very appealing properties. Most notably it is simple, Pareto efficient, and strategy-proof (i.e., it is immune to misrepresentation of preferences). Unfortunately this model cannot capture an important feature present in many real-life applications: the existence of tenants who already live in a house and who can keep on doing so. For example, professors are usually entitled to keep their current offices and students in many campuses can keep their on-campus houses up to three or four years. There is no reason to think that these agents may not wish to move to another house (or a new office in case of a professor). Often in practice, those who want to move are asked to give up their houses before they are assigned another one. Since there are no guarantees of getting a better house (or whether they will get a house at all), many of them simply keep their current houses, which results in loss of potential gains from trade.<sup>2</sup> As an implication, in many cases the outcome of the random serial dictatorship (or even its version that favors the existing tenants) is not Pareto efficient. This observation is our main motivation.

We introduce a richer class of problems that includes existing tenants as well as new applicants and propose a class of mechanisms, namely *top trading cycles mechanisms*, which are not only Pareto efficient, but also individually rational and strategy-proof. By individual rationality all agents are assured a house that is at least as good as their own and therefore there is no risk in applying for a new house. By strategy-proofness truth-telling is a dominant strategy so there is no point in misrepresenting preferences. Once full participation and sincere revelation are guaranteed, Pareto

<sup>2</sup> Clearly the same problem may occur in any model where agents have outside options. See for example Krishna and Perry [10] for a discussion of this issue in an abstract setup.

efficiency of these mechanisms ensures the Pareto efficiency of the final outcome. Top trading cycles mechanisms have two additional features that are crucial for practical purposes: (i) they are simple *direct mechanisms*: agents report their preferences and the outcome is obtained using one of the two algorithms that we provide; and (ii) they can accommodate any hierarchy of seniorities (as in the case of random serial dictatorship).

This paper, to our knowledge, is the first to analyze a house allocation model where there are both existing tenants and new applicants.<sup>3</sup> The model with only new applicants is introduced by Hylland and Zeckhauser [8]<sup>4</sup> and the model with only existing tenants is introduced by Shapley and Scarf [20].<sup>5</sup> The *top trading cycles mechanism from random orderings* reduces to random serial dictatorship in the first model and to the competitive mechanism in the second, which are the predominant mechanisms in these models. Therefore it integrates ideas from both literatures.<sup>6</sup>

The organization of the rest of the paper is as follows: In Section 2 we motivate our proposed mechanism by restricting attention to the simple case with one existing tenant. In Section 3 we give the formal model and definitions. We devote Section 4 to real-life mechanisms. In Section 5 we introduce and study the class of top trading cycles mechanisms. In Section 6 we provide extensions that are useful for real-life applications. Finally, in the Appendix, we present an omitted proof.

## 2. PREVIEW

In this section, we motivate the paper and give an intuition for our proposed mechanism in the simple case of one existing tenant.

A number of houses has to be allocated to a group of agents through a lottery. The agents are randomly ordered and the first agent is assigned his or her top choice, the next agent his or her top choice among the remaining houses, and so on. One of the houses is already occupied and its tenant is given two options: He or she can keep that house or give it up and enter

<sup>3</sup> See Thomson [22] for an exchange economy where agents have individual endowments as well as a collective endowment.

<sup>4</sup> See also Abdulkadiroğlu and Sönmez [1], Bogomolnaia and Moulin [2], Collins and Krishna [3], Papai [13], Svensson [21], and Zhou [23].

<sup>5</sup> See also Garratt and Qin [5], Ma [11], Roth [14], and Roth and Postlewaite [18].

<sup>6</sup> The mechanism design approach to matching problems has been especially fruitful in the context of *two-sided matching markets*. Roth [15] shows that the *hospital-proposing deferred acceptance mechanism* (Gale and Shapley [4]) has been in use in the U.S. hospital-intern market since 1950. It is well known that, this mechanism favors hospitals and starting with the 1998 match an alternative mechanism that favors the interns will be used (Roth and Peranson [17]). See also Roth [16] and Mongell and Roth [12] for case studies concerning British hospital-intern markets and admissions to American sororities, respectively.

the lottery. Since there is no guarantee that he or she will get a better house, the existing tenant may choose the first option, which in turn may result in a loss of potential gains from trade.

EXAMPLE 1. There are three agents  $i_1, i_2, i_3$ , and three houses  $h_1, h_2, h_3$ . Agent  $i_1$  is a current tenant and occupies house  $h_1$ . Agents  $i_2, i_3$  are new applicants and houses  $h_2, h_3$  are vacant houses. The following matrix gives the utilities of agents over houses:

	$h_1$	$h_2$	$h_3$
$i_1$	3	4	1
$i_2$	4	3	1
$i_3$	3	4	1

Agent  $i_1$  has two options:

1. keep house  $h_1$  or
2. give it up and enter the lottery.

Utility from keeping house  $h_1$  is 3. If he or she enters the lottery then there are several possibilities depending on the chosen ordering, summarized in the following table:

ordering	assignment of $i_1$	assignment of $i_2$	assignment of $i_3$
$i_1-i_2-i_3$	$h_2$	$h_1$	$h_3$
$i_1-i_3-i_2$	$h_2$	$h_3$	$h_1$
$i_2-i_1-i_3$	$h_2$	$h_1$	$h_3$
$i_2-i_3-i_1$	$h_3$	$h_1$	$h_2$
$i_3-i_1-i_2$	$h_1$	$h_3$	$h_2$
$i_3-i_2-i_1$	$h_3$	$h_1$	$h_2$

Assuming that agent  $i_1$  is an expected-utility maximizer, utility from entering the lottery is

$$\frac{1}{6} u(h_1) + \frac{3}{6} u(h_2) + \frac{2}{6} u(h_3) = \frac{3}{6} + \frac{12}{6} + \frac{2}{6} = \frac{17}{6}.$$

Therefore the optimal strategy is keeping house  $h_1$ . Since both agent  $i_2$  and agent  $i_3$  prefer house  $h_2$  to house  $h_3$ , the eventual outcome is

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ h_1 & h_2 & h_3 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} i_1 & i_2 & i_3 \\ h_1 & h_3 & h_2 \end{pmatrix},$$

both with  $1/2$  probability. Among these two matchings the first is Pareto dominated by

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ h_2 & h_1 & h_3 \end{pmatrix}$$

and therefore this mechanism may lead to Pareto inefficient outcomes.

### 2.1. *Avoiding Inefficiency with One Existing Tenant*

The cause for the inefficiency is that the mechanism fails to guarantee the existing tenant a house that is at least as good as the one he or she already holds. Therefore we should fix this "deficiency." One natural modification that will do the trick in the simple case with one existing tenant is the following:

1. Order the agents by means of a lottery.
2. Assign the first agent his or her top choice, the second agent his or her top choice among the remaining houses, and so on, *until someone demands the house the existing tenant holds.*
3. (a) If the existing tenant is already assigned a house, then do not disturb the procedure.  
(b) If the existing tenant is not assigned a house, then modify the remainder of the ordering by inserting him or her at the top, and proceed with the procedure.

In this way the existing tenant cannot lose his or her house without getting a better one, and therefore has nothing to loose from participating. We illustrate the modified mechanism with Example 1.

EXAMPLE 1 (Continued). The lottery can result in six orderings. If the ordering is one of  $i_1-i_2-i_3$ ,  $i_1-i_3-i_2$ , or  $i_3-i_1-i_2$ , then agent  $i_1$  leaves before anyone demands house  $h_1$  and therefore the resulting allocation is not affected:

initial ordering	modified ordering	assignment of $i_1$	assignment of $i_2$	assignment of $i_3$
$i_1-i_2-i_3$	$i_1-i_2-i_3$	$h_2$	$h_1$	$h_3$
$i_1-i_3-i_2$	$i_1-i_3-i_2$	$h_2$	$h_3$	$h_1$
$i_3-i_1-i_2$	$i_3-i_1-i_2$	$h_1$	$h_3$	$h_2$

If the ordering is  $i_2-i_1-i_3$  or  $i_2-i_3-i_1$ , then in the first step agent  $i_2$  demands house  $h_1$ . In both cases the ordering is changed to  $i_1-i_2-i_3$  and the resulting outcome is as follows:

initial ordering	modified ordering	assignment of $i_1$	assignment of $i_2$	assignment of $i_3$
$i_2-i_1-i_3$	$i_1-i_2-i_3$	$h_2$	$h_1$	$h_3$
$i_2-i_3-i_1$	$i_1-i_2-i_3$	$h_2$	$h_1$	$h_3$

Finally if the ordering is  $i_3-i_2-i_1$ , then in the first step agent  $i_3$  is assigned house  $h_2$  and in the next step agent  $i_2$  demands house  $h_1$ . The remainder of the ordering is changed to  $i_1-i_2$  and this results in the following outcome:

initial ordering	modified ordering	assignment of $i_1$	assignment of $i_2$	assignment of $i_3$
$i_3-i_2-i_1$	$i_3-i_1-i_2$	$h_1$	$h_3$	$h_2$

Therefore the modified mechanism selects one of

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ h_2 & h_1 & h_3 \end{pmatrix}, \begin{pmatrix} i_1 & i_2 & i_3 \\ h_2 & h_3 & h_1 \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} i_1 & i_2 & i_3 \\ h_1 & h_3 & h_2 \end{pmatrix},$$

with probabilities of  $1/2$ ,  $1/6$ , and  $1/3$  respectively. Note that all these matchings are Pareto efficient.

### 3. THE MODEL

We are now ready to introduce the formal model. *A house allocation problem with existing tenants*, or simply a *problem*, 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 preference relations  $P = (P_i)_{i \in I_E \cup I_N}$ .

Let  $I = I_E \cup I_N$  denote the set of all agents and  $H = H_O \cup H_V \cup \{h_0\}$  denote the set of all houses plus the null house. (Here the null house  $h_0$  denotes the no house option.) Every existing tenant  $i \in I_E$  is endowed with (i.e., currently lives in) the occupied house  $h_i \in H_O$ . Every agent  $i \in I$  has a strict preference relation  $P_i$  on  $H$ . Let  $R_i$  denote the at-least-as-good-as relation associated with  $P_i$ . We assume that the null house  $h_0$  is the last choice for each agent. This assumption is for expositional simplicity and it

is not essential. We later discuss the modifications needed in the absence of this assumption. Let  $\mathcal{P}$  be the class of all such preferences on  $H$ .

A (house) matching  $\mu$  is an assignment of houses to agents such that

1. every agent is assigned one house, and
2. only the null house  $h_0$  can be assigned to more than one agent.

To say that an agent is assigned the null house is to say the agent is not assigned any “real” house at all. For any agent  $i \in I$ , we refer to  $\mu(i)$  as the assignment of agent  $i$  under  $\mu$ . Let  $\mathcal{M}$  be the set of all matchings. Given a preference relation  $P_i$  of an agent  $i$ , initially defined over  $H$ , we extend it to the set of matchings in the following natural way: Agent  $i$  prefers the matching  $\mu$  to matching  $\nu$  if and only if he prefers  $\mu(i)$  to  $\nu(i)$ . That is, our model is one with no consumption externalities. A matching is *Pareto efficient* if there is no other matching that makes all agents weakly better off and at least one agent strictly better off. A matching is *individually rational* if no existing tenant strictly prefers his endowment to his assignment.

A housing lottery is a probability distribution over all matchings. Let  $\Delta\mathcal{M}$  denote the set of all housing lotteries. A housing lottery is *ex post Pareto efficient* if it gives positive weight to only Pareto efficient matchings.

A matching mechanism consists of

1. a strategy space  $S_i$  for every agent  $i \in I$ , and
2. an outcome function  $\varphi: \prod_{i \in I} S_i \rightarrow \mathcal{M}$  which assigns a matching for every strategy-profile.

Similarly a lottery mechanism consists of

1. a strategy space  $S_i$  for every agent  $i \in I$ , and
2. an outcome function  $\psi: \prod_{i \in I} S_i \rightarrow \Delta\mathcal{M}$  which assigns a housing lottery for every strategy-profile.

A direct matching mechanism is a matching mechanism where the strategy space is the set of all strict preferences  $\mathcal{P}$  for all agents. A direct lottery mechanism is defined analogously. A direct matching mechanism is *individually rational* if it always selects individually rational matchings and it is *Pareto efficient* if it always selects Pareto efficient matchings. A direct lottery mechanism is *ex post individually rational* if it gives positive weight to only individually rational matchings and it is *ex post Pareto efficient* if it gives positive weight to only Pareto efficient matchings. A direct matching or lottery mechanism is *strategy-proof* (or *dominant strategy incentive compatible*) if truth-telling is a dominant strategy in its associated preference revelation game.

## 4. REAL-LIFE MECHANISMS

The *random serial dictatorship* and its variants are commonly used in real-life applications. Some examples include graduate housing at Michigan, Princeton, Rochester, and Stanford and undergraduate housing at Carnegie–Mellon, Duke, Harvard, Michigan, Northwestern, Pennsylvania, and Yale. We need the following additional definitions in order to define this mechanism and some of its common variants.

Given a strict preference  $P_i$ , the *choice* of agent  $i$  from a set of houses  $G \subseteq H$  is the best house in  $G$ . Given a group  $J \subseteq I$  of agents, an *ordering* of these agents is a one-to-one function  $f: \{1, \dots, |J|\} \rightarrow J$ . Here agent  $f(1)$  is ordered first, agent  $f(2)$  is ordered second, and so on. Given a group  $J \subseteq I$  of agents and a set  $G \subseteq H$  of houses, the *serial dictatorship* induced by ordering  $f$  is defined as follows: The agent who is ordered first under  $f$  gets his or her top choice from  $G$ , the next agent gets his or her top choice among the remaining houses, and so on. Since the preferences are strict, the serial dictatorship assigns every agent in  $J$  a unique house (possibly the null house). The *random serial dictatorship* randomly selects an ordering from a given distribution of orderings<sup>7</sup> and uses the induced serial dictatorship.

4.1. *Random Serial Dictatorship with Squatting Rights*

We are now ready to introduce a variant of the random serial dictatorship that allows the existing tenants to keep their current houses.<sup>8</sup> This version is used for undergraduate housing at Carnegie–Mellon, Duke, Harvard, Northwestern, and Pennsylvania, among others. The strategy spaces and the outcome function for the *random serial dictatorship with squatting rights* are as follows:

*Strategy spaces.*

1. *Existing tenants:* Every existing tenant  $i \in I_E$  announces whether he or she is In or Out and a strict preference  $Q_i$  over all “real” houses.<sup>9</sup> Formally his or her strategy space is  $S_i = S_{i1} \times S_{i2} = \{\text{In}, \text{Out}\} \times \mathcal{P}$ .

2. *New applicants:* Every new applicant  $i \in I_N$  announces a strict preference  $Q_i$  over all houses. Formally his or her strategy space is  $S_i = \mathcal{P}$ .

<sup>7</sup> This distribution varies widely from case to case. While it is uniform in some cases, it is dictated by seniority in others. There may be variations even within a campus. For example at Yale the chosen ordering is completely random in some residential colleges and it is dictated by seniority in others.

<sup>8</sup> This practice is called “squatting.”

<sup>9</sup> In most real-life applications, existing tenants announce “Out” by simply not applying to move to another house.



*Outcome function.* For a given strategy profile  $s \in \prod_{i \in I} S_i$ , the housing lottery is obtained as follows:

1. Every existing tenant  $i \in I_E$  who announces Out is assigned his or her current house  $h_i$  with certainty.

2. Let  $J = I_N \cup \{i \in I_E : s_{iI} = \text{In}\}$  and  $G = H_V \cup \{h_i \in H_O : s_{iI} = \text{In}\}$ . That is,  $J$  consists of the set of new applicants together with the existing tenants who announce In and  $G$  consists of vacant houses plus the houses of existing tenants who announce In.

(a) An ordering  $f$  of agents in  $J$  is randomly chosen from a given distribution of orderings.

(b) Houses in  $G$  are assigned to these agents based on the serial dictatorship induced by  $f$  under the announced preferences.

In everyday language, those existing tenants who want to move to another house need to give up their houses and join the group of new applicants while their houses are added to the pool of vacant houses.

In many cases (for example at Duke and some of the residences at Pennsylvania) the lottery favors the existing tenants over the newcomers. In such cases the mechanism is implemented as follows: Existing tenants who want to move give up their houses. These houses are added to the pool of vacant houses. Existing tenants who give up their houses are randomly ordered with a lottery and the induced serial dictatorship is used to determine their assignments. Next an ordering of the new applicants is randomly chosen with another lottery and the remaining houses are allocated among the new applicants based on the induced serial dictatorship. In many cases this second step is carried out several months after the first one.

While the *random serial dictatorship with squatting rights* is very popular, it suffers a major deficiency: Since it does not guarantee each existing tenant a house that is at least as good as his own, some of them may choose to stay Out (i.e., use their squatting rights), and this may result in the loss of potentially large gains from trade. So in most cases the eventual house allocation is Pareto inefficient.<sup>10</sup>

We could find only two mechanisms which gives the existing tenants the right to keep their current houses without imposing them to opt out. They are the *random serial dictatorship with waiting list* and the *MIT-NH4 mechanism*.

<sup>10</sup> If on the other hand there are no existing tenants (and thus no occupied house), the random serial dictatorship is Pareto efficient. Moreover it is equivalent to *core from random endowments* which randomly selects an initial allocation and finds the unique core allocation of the induced market (Abdulkadiroğlu and Sönmez [1], Knuth [9]).

#### 4.2. Random Serial Dictatorship with Waiting List

The *random serial dictatorship with waiting list* is used for graduate housing at Rochester. It is a direct mechanism and thus agents announce their preferences over all houses. For a given ordering  $f$  of agents, the outcome is obtained as follows:

Define the set of *available* houses for Step 1 to be the set of vacant houses. Define the set of *acceptable* houses for agent  $i$  to be the set of all houses in case agent  $i$  is a new applicant, and the set of all houses better than his or her current house  $h_i$  in case he or she is an existing tenant.

*Step 1.* The agent with the highest priority among those who have at least one acceptable available house is assigned his or her top available house and removed from the process. His or her assignment is deleted from the set of available houses for Step 2. In case he or she is an existing tenant, his or her current house becomes available for Step 2. If there is at least one remaining agent and one available house that is acceptable to at least one of them, then we go to the next step.

⋮

*Step  $t$ .* The set of available houses for Step  $t$  is defined at the end of Step  $t - 1$ . The agent with the highest priority among all remaining agents who has at least one acceptable available house is assigned his or her top available house and removed from the process. His or her assignment is deleted from the set of available houses for Step  $t + 1$ . In case he or she is an existing tenant, his or her current house becomes available for Step  $t + 1$ . If there is at least one remaining agent and one available house that is acceptable to at least one of them, then we go to the next step.

When the process terminates, those existing tenants who are not re-assigned keep their current houses.

This variant of the random serial dictatorship is also inefficient. Consider the following example.

**EXAMPLE 2.** Let  $I_E = \{i_1, i_2, i_3\}$ ,  $I_N = \emptyset$ ,  $H_O = \{h_1, h_2, h_3\}$ , and  $H_V = \{h_4\}$ . Here the existing tenant  $i_k$  occupies the house  $h_k$  for  $k = 1, 2, 3$ . Let the agents be ordered as  $i_1 - i_2 - i_3$  and let the preferences (from best to worst) be as follows:

$P_{i_1}$	$P_{i_2}$	$P_{i_3}$
$h_2$	$h_3$	$h_1$
$h_3$	$h_1$	$h_4$
$h_1$	$h_2$	$h_3$
$h_4$	$h_4$	$h_2$
$h_0$	$h_0$	$h_0$

Let us find the outcome of the *random serial dictatorship with waiting list*:

*Step 1.* The only available house at Step 1 is house  $h_4$ . It is acceptable to only agent  $i_3$ . So, agent  $i_3$  is assigned house  $h_4$ .

*Step 2.* The only available house at Step 2 is house  $h_3$ . It is acceptable to both agent  $i_1$  and agent  $i_2$ . Since agent  $i_1$  has the higher priority, agent  $i_1$  is assigned house  $h_3$ .

*Step 3.* The only available house at Step 3 is house  $h_1$ . It is acceptable to agent  $i_2$ . So agent  $i_2$  is assigned house  $h_1$ .

Since there are no remaining agents at the end of Step 3, the process terminates and the final matching is

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ h_3 & h_1 & h_4 \end{pmatrix}$$

which is Pareto dominated by

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ h_2 & h_3 & h_1 \end{pmatrix}$$

#### 4.3. MIT-NH4 Mechanism

The following mechanism is used at the residence NH4 of MIT. It works as follows:

1. An ordering  $f$  of agents is chosen from a given distribution of agents.
2. The first agent is *tentatively* assigned his or her top choice among all houses, the next agent is *tentatively* assigned his top choice among the remaining houses, and so on, until a *squatting conflict* occurs.
3. A *squatting conflict* occurs if it is the turn of an existing tenant but every remaining house is worse than his or her current house. That means someone else, the *conflicting agent*, is tentatively assigned the existing tenant's current house. When this happens

(a) the existing tenant is assigned his or her current house and removed from the process, and

(b) all tentative assignments starting with the conflicting agent and up to the existing tenant are erased.

At this point the squatting conflict is resolved and the process starts over again with the conflicting agent. Every squatting conflict that occurs afterwards is resolved in a similar way.

4. The process is over when there are no houses or agents left. At this point all tentative assignments are finalized.

While it is innovative, the *MIT-NH4 mechanism* does not resolve the inefficiency problem.<sup>11</sup>

EXAMPLE 3. Let  $I_E = \{i_1, i_2, i_3, i_4\}$ ,  $I_N = \{i_5\}$ ,  $H_O = \{h_1, h_2, h_3, h_4\}$ , and  $H_V = \{h_5\}$ . Here the existing tenant  $i_k$  occupies the house  $h_k$  for  $k = 1, 2, 3, 4$ . Let the ordering  $f$  order the agents as  $i_1-i_2-i_3-i_4-i_5$  and let the preferences be as follows:

$P_{i_1}$	$P_{i_2}$	$P_{i_3}$	$P_{i_4}$	$P_{i_5}$
$h_3$	$h_4$	$h_5$	$h_3$	$h_4$
$h_4$	$h_5$	$h_3$	$h_5$	$h_5$
$h_5$	$h_2$	$h_4$	$h_4$	$h_3$
$h_1$	$h_3$	$h_2$	$h_2$	$h_1$
$h_2$	$h_1$	$h_1$	$h_1$	$h_2$
$h_0$	$h_0$	$h_0$	$h_0$	$h_0$

Let us find the outcome of the MIT-NH4 mechanism:

1. First agent  $i_1$  is tentatively assigned  $h_3$ , next agent  $i_2$  is tentatively assigned  $h_4$ , next agent  $i_3$  is tentatively assigned  $h_5$ , and next its agent  $i_4$ 's turn and a squatting conflict occurs. The conflicting agent is agent  $i_2$  who was tentatively assigned  $h_4$ . Agent  $i_2$ 's tentative assignment, as well as that of agent  $i_3$ , is erased. Agent  $i_4$  is assigned his or her current house  $h_4$  and removed from the process. This resolves the squatting conflict.

2. The process starts over with the conflicting agent  $i_2$ . Agent  $i_2$  is tentatively assigned  $h_5$  and next it is agent  $i_3$ 's turn and another squatting conflict occurs. The conflicting agent is agent  $i_1$  who was tentatively assigned  $h_3$ . His tentative assignment, as well as that of agent  $i_2$  are erased. Agent  $i_3$  is assigned his current house  $h_3$  and removed from the process. This resolves the second squatting conflict.

3. The process starts over with the conflicting agent  $i_1$ . He is tentatively assigned  $h_5$ , next agent  $i_2$  is tentatively assigned  $h_2$  and finally agent  $i_5$  is tentatively assigned  $h_1$ . At this point all tentative assignments are finalized.

Therefore the final matching is

$$\begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 \\ h_5 & h_2 & h_3 & h_4 & h_1 \end{pmatrix}$$

<sup>11</sup> The housing rules at NH4 allows two agents to swap their assignments, but only if they have the consent of every agent whose ordering is between them.



points to and removed with his assignment. If there is an available house in a cycle then the agent with the highest priority in this step is also in the same cycle. If this agent is an existing tenant, then his or her house cannot be in any cycle and it becomes available for Step  $t + 1$ . All available houses that are not removed remain available. If there is at least one remaining agent and one remaining house then we go to the next step.

By the finiteness of the numbers of agents and houses at least one cycle forms at each step. Hence this algorithm terminates in at most  $\min\{|I|, |H|\}$  steps. Any agent who is not assigned a house at the termination is assigned the null house  $h_0$  (i.e., remains unmatched).

*Remark 1.* We indicated that at each step of the top trading cycles algorithm there is at least one cycle and each of them should be removed. Nevertheless, if one removes one cycle at a step, the outcome is still the same. This is because a cycle that is not removed at any step remains a cycle at the next step.

When there are no new applicants and no vacant houses this algorithm reduces to the celebrated *Gale's top trading cycles algorithm*.<sup>13</sup> Our algorithm is very much inspired by Gale's algorithm. On the other hand, when there are no existing tenants it reduces to the serial dictatorship induced by  $f$ .

### 5.1. An Example

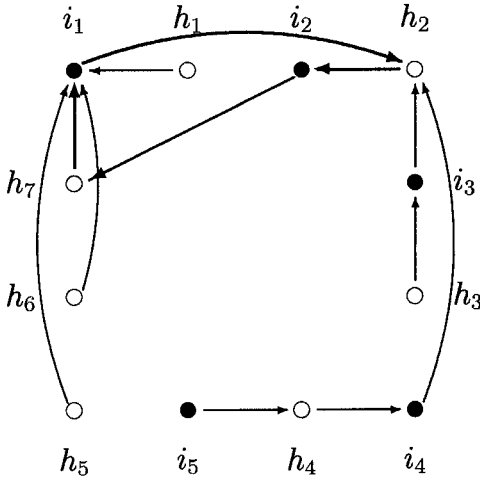
In this section we give a detailed example in order to illustrate the dynamics of the top trading cycles algorithm.

Let  $I_E = \{i_1, i_2, i_3, i_4\}$ ,  $I_N = \{i_5\}$ ,  $H_O = \{h_1, h_2, h_3, h_4\}$ , and  $H_\emptyset = \{h_5, h_6, h_7\}$ . Here the existing tenant  $i_k$  occupies the house  $h_k$  for  $k = 1, \dots, 4$ . Let the ordering  $f$  order the agents as  $i_1 - i_2 - i_3 - i_4 - i_5$  and let the preferences be as follows:

$P_{i_1}$	$P_{i_2}$	$P_{i_3}$	$P_{i_4}$	$P_{i_5}$
$h_2$	$h_7$	$h_2$	$h_2$	$h_4$
$h_6$	$h_1$	$h_1$	$h_4$	$h_3$
$h_5$	$h_6$	$h_4$	$h_3$	$h_7$
$h_1$	$h_5$	$h_7$	$h_6$	$h_1$
$h_4$	$h_4$	$h_3$	$h_1$	$h_2$
$h_3$	$h_3$	$h_6$	$h_7$	$h_5$
$h_7$	$h_2$	$h_5$	$h_5$	$h_6$
$h_0$	$h_0$	$h_0$	$h_0$	$h_0$

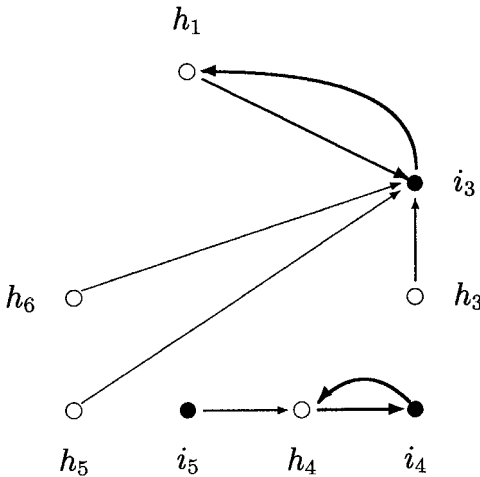
<sup>13</sup> In this case our model reduces to *housing markets* introduced by Shapley and Scarf [20]. In this context Roth and Postlewaite [18] shows that there is a unique core allocation for every market which also coincides with the unique competitive allocation. Gale's top trading cycles algorithm can be used to find this allocation.

Step 1.



The set of available houses in Step 1 is  $H_V = \{h_5, h_6, h_7\}$ . The only cycle that is formed at this step is  $(i_1, h_2, i_2, h_7)$ . Therefore  $i_1$  is assigned  $h_2$  and  $i_2$  is assigned  $h_7$ .

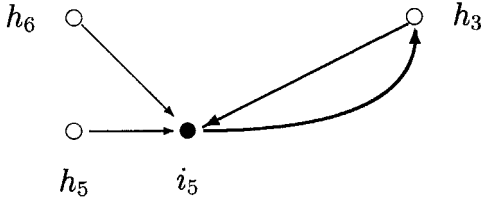
Step 2.



Since agent  $i_1$  leaves in Step 1, house  $h_1$  becomes available in Step 2. Therefore the set of available houses for Step 2 is  $\{h_1, h_5, h_6\}$ . The available houses  $h_1, h_5$ , and  $h_6$  all point to agent  $i_3$ , now the highest ranking agent.

There are two cycles  $(i_3, h_1)$  and  $(i_4, h_4)$ . Therefore  $i_3$  is assigned  $h_1$  and  $i_4$  is assigned his or her own house  $h_4$ .

*Step 4.*



Since  $i_3$  leaves in Step 2, house  $h_3$  becomes available for Step 3. Therefore the set of available houses for Step 3 is  $\{h_3, h_5, h_6\}$ . The available houses  $h_3, h_5$ , and  $h_6$  all point to the only remaining agent  $i_5$ . The only cycle is  $(i_5, h_3)$ . Therefore  $i_5$  is assigned  $h_3$ . There are no remaining agents so the algorithm terminates and the matching it induces is:

$$\begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 \\ h_2 & h_7 & h_1 & h_4 & h_3 \end{pmatrix}$$

## 5.2. Efficiency, Individual Rationality, and Strategy-Proofness

For any ordering  $f$ , the induced *top trading cycles mechanism*  $\varphi^f$  has some very desirable properties. In particular it is

1. Pareto efficient,
2. individually rational, and
3. strategy proof.

The significance of the first two properties is clear. Strategy-proofness is also crucial since in its absence agents may attempt to manipulate by reporting fake preferences. If they do, the resulting matching is efficient under announced preferences but not necessarily under the true preferences. We next prove these properties.

**PROPOSITION 1.** *For any ordering  $f$ , the induced top trading cycles mechanism  $\varphi^f$  is Pareto efficient.*

*Proof.* Consider the top trading cycles algorithm. Any agent who leaves at Step 1 is assigned his or her top choice and cannot be made better off. Any agent who leaves at Step 2 is assigned his or her top choice among those houses remaining at Step 2 and since the preferences are strict he or



she cannot be made better off without hurting someone who left at Step 1. Proceeding in a similar way, no agent can be made better off without hurting someone who left at an earlier step. Therefore the mechanism  $\varphi^f$  is Pareto efficient. ■

*Remark 2.* The criticism of the mechanisms described in Section 4 is not that their outcomes are necessarily Pareto inefficient, but that they can be Pareto inefficient at some preference profiles for some orderings of the agents. The top trading cycles mechanism, in contrast, has the advantage that its outcome is Pareto efficient at every preference profile regardless of the ordering of the agents.

**PROPOSITION 2.** *For any ordering  $f$ , the induced top trading cycles mechanism  $\varphi^f$  is individually rational.*

*Proof.* Consider the top trading cycles algorithm. For any existing tenant  $i \in I_E$ , his or her house  $h_i$  points to him or her until he or she leaves. Therefore the assignment of  $i$  cannot be worse than his endowment  $h_i$ . ■

Our next result is a generalization of a similar result in Roth [14] who shows strategy-proofness for the special case of housing markets.<sup>14</sup> The following lemma is the key to our proof.

**LEMMA 1.** *Fix the announced preferences of all agents except  $i$  at  $Q_{-i} = (Q_j)_{j \in I \setminus \{i\}}$ . Suppose that in the top trading cycles algorithm agent  $i$  leaves at Step  $T$  under  $Q_i$  and at Step  $T^*$  under  $Q_i^*$ . Suppose  $T \leq T^*$ . Then the remaining agents and houses at the beginning of Step  $T$  are the same whether agent  $i$  announces  $Q_i$  or  $Q_i^*$ .*

*Proof.* Since agent  $i$  fails to participate in a cycle prior to Step  $T$  in either case, the same cycles form and therefore the same agents and houses leave before Step  $T$ . ■

**THEOREM 1.** *For any ordering  $f$ , the induced top trading cycles mechanism  $\varphi^f$  is strategy proof.*

*Proof.* Consider an agent  $i$  with true preferences  $P_i$ . Fix an announced preference profile  $Q_{-i} = (Q_j)_{j \in I \setminus \{i\}}$  for every agent except  $i$ . We want to show that revealing his or her true preferences  $P_i$  is at least as good as announcing any other preferences  $Q_i$ . Let  $T$  be the step at which agent  $i$  leaves under  $Q_i$ , let  $(h, j_1, j_2, \dots, j_k, i)$  be the cycle  $i$  joins, and thus

<sup>14</sup> Roth uses graph theoretic techniques in his proof which can be used to prove Theorem 1 as well. We, on the other hand, provide a game theoretic proof.

house  $h$  be  $i$ 's assignment. Let  $T^*$  be the step at which  $i$  leaves under his or her true preferences  $P_i$ . We want to show that  $i$ 's assignment under  $P_i$  is at least as good as house  $h$ . We have two cases to consider.

*Case 1.*  $T^* \geq T$ .

Suppose agent  $i$  announces his or her true preferences  $P_i$ . Consider Step  $T$ . By Lemma 1, the same agents and houses remain in the market at the beginning of this step whether agent  $i$  announces  $Q_i$  or  $P_i$ . Therefore at Step  $T$ , house  $h$  points to agent  $j_1$ , agent  $j_1$  points to house  $j_2$ , ..., house  $j_k$  points to agent  $i$ . Moreover, they keep doing so as long as agent  $i$  remains. Since agent  $i$  truthfully points to his or her best choice at each step,  $i$  either receives a house that is at least as good as  $h$  or eventually joins the cycle  $(h, j_1, j_2, \dots, j_k, i)$  and receives house  $h$ .

*Case 2.*  $T^* < T$ .

By Lemma 1 the same houses remain in the market at the beginning of Step  $T^*$  whether agent  $i$  announces  $Q_i$  or  $P_i$ . Moreover, agent  $i$  is assigned his or her best choice remaining at Step  $T^*$  under  $P_i$ . Therefore, in this case too  $i$ 's assignment under the true preferences  $P_i$  is at least as good as house  $h$ . ■

### 5.3. Respecting Seniority

There are many applications where agents are naturally ordered based on their seniority. Let  $f$  denote this ordering. We earlier suggested that in such applications one can use the top trading cycles mechanism  $\varphi^f$ . We next justify this suggestion.

In order to simplify the exposition we refer to any mechanism that is Pareto efficient, individually rational, and strategy-proof as an *admissible* mechanism. Before committing to mechanism  $\varphi^f$ , one should ask the following question: Is there another admissible mechanism that always better respects the seniority of the agents? The answer to this question is negative. Suppose that for some preference profile an admissible mechanism  $\varphi$  assigns an agent a better house than  $\varphi^f$  does. Then one can construct a preference profile in which an agent with higher seniority is worse off under  $\varphi$  than he or she is under  $\varphi^f$ .

**THEOREM 2.** *Let  $\varphi$  be a mechanism that is Pareto efficient, individually rational, and strategy-proof. Then for all  $t \in \{1, \dots, |I|\}$ ,*

$$\begin{aligned} & [\exists P \in \mathcal{P} \text{ such that } \varphi_{f(t)}(P) P_{f(t)} \varphi_{f(t)}^f(P)] \\ & \Rightarrow [\exists \tilde{P} \in \mathcal{P} \text{ and } s < t \text{ such that } \varphi_{f(s)}^f(\tilde{P}) \tilde{P}_{f(s)} \varphi_{f(s)}(\tilde{P})]. \end{aligned}$$

(See the Appendix for a proof.)

An alternative way to read Theorem 2 is as follows: As far as agent  $f(1)$  is concerned, the mechanism  $\varphi^f$  assigns him a house that is at least as good as the assignment of any admissible mechanism at all preference profiles. Next consider all admissible mechanisms that perform equally well for agent  $f(1)$ . (Equivalently consider all admissible mechanisms that agree with mechanism  $\varphi^f$  on the assignment of agent  $f(1)$  at every preference profile.) The mechanism  $\varphi^f$  assigns agent  $f(2)$  a house that is at least as good as the assignment of any such mechanism at all preference profiles. In general, consider all admissible mechanisms that perform equally well for agents  $f(1), f(2), \dots, f(k)$  where  $k < |I|$ . The mechanism  $\varphi^f$  assigns agent  $f(k+1)$  a house that is at least as good as the assignment of any such mechanism at all preference profiles.

#### 5.4. An Alternative Algorithm

We can find the outcome of the top trading cycles mechanism using the following *you request my house—I get your turn* (or in short YRMH-IGYT) algorithm: For any given ordering  $f$ , assign the first agent his or her top choice, the second agent his or her top choice among the remaining houses, and so on, until someone demands the house of an existing tenant. If at that point the existing tenant whose house is demanded is already assigned a house, then do not disturb the procedure. Otherwise modify the remainder of the ordering by inserting him to the top and proceed. Similarly, insert any existing tenant who is not already served at the top of the line once his or her house is demanded. If at any point a loop forms, it is formed by exclusively existing tenants and each of them demands the house of the tenant next in the loop. (A *loop* is an ordered list of agents  $(i_1, i_2, \dots, i_k)$  where agent  $i_1$  demands the house of agent  $i_2$ , agent  $i_2$  demands the house of agent  $i_3, \dots$ , agent  $i_k$  demands the house of agent  $i_1$ .) In such cases remove all agents in the loop by assigning them the houses they demand and proceed.

Note that if there is only one existing tenant, YRMH-IGYT algorithm reduces to the algorithm we defined in Section 2.1.

**THEOREM 3.** *For a given ordering  $f$ , the YRMH-IGYT algorithm yields the same outcome as the top trading cycles algorithm.*

*Proof.* For any set  $J$  of agents and set  $G$  of houses remaining in the algorithm, YRMH-IGYT algorithm assigns the next series of houses in one of two possible ways.

*Case 1.* There is a sequence of agents  $i_1, i_2, \dots, i_k$  (which may consist of a single agent) where agent  $i_1$  has the highest priority in  $J$  and demands

house of  $i_2$ , agent  $i_2$  demands house of  $i_3$ , ..., agent  $i_{k-1}$  demands house of  $i_k$ , and  $i_k$  demands an available house  $h$ . At this point agent  $i_k$  is assigned house  $h$ , the next agent  $i_{k-1}$  is assigned house  $h_{i_k}$  (which just became available), ..., and finally agent  $i_1$  is assigned house  $h_{i_2}$ . Note that the ordered list  $(h, i_1, h_{i_2}, i_2, \dots, h_{i_k}, i_k)$  is a (top trading) cycle for the pair  $(J, G)$ .

*Case 2.* There is a loop  $(i_1, i_2, \dots, i_k)$  of agents. When that happens agent  $i_1$  is assigned the house of  $i_2$ , agent  $i_2$  is assigned house of  $i_3$ , ..., agent  $i_k$  is assigned house of  $i_1$ . In this case  $(h_{i_1}, i_1, h_{i_2}, i_2, \dots, h_{i_k}, i_k)$  is a (top trading) cycle for the pair  $(J, G)$ .

Hence the YRMH-IGYT algorithm locates a cycle and implements the associated trades for any sets of remaining agents and houses. This observation together with Remark 1 implies the desired result. ■

The two algorithms have different merits and hence we consider both of them to be useful: It is easier to see how the top trading cycles algorithm works and its use makes the results in the previous sections easier to prove. The YRMH-IGYT algorithm, on the other hand, has an intuitive interpretation and also it can be coded on computer as it is.

### 5.5. Top Trading Cycles Mechanisms from Random Orderings

Next we propose a class of lottery mechanisms: Randomly choose an ordering from any exogenous distribution of orderings and use the induced top trading cycles mechanism. We refer to this class as *top trading cycles mechanisms from random orderings*. In some applications one may want to favor a group of agents (such as existing tenants) over another (such as new applicants). In other applications, favoring one group or another may not be appropriate. This class of lottery mechanisms is rich enough to accommodate any form of priorities. Moreover, since orderings are drawn from exogenous distributions, they are ex post individually rational, ex post Pareto efficient, and strategy proof. We state this observation as a corollary.

**COROLLARY 1.** *Any top trading cycles mechanism from random orderings is ex post individually rational, ex post Pareto efficient, and strategy-proof.*

*Proof.* Immediate from Propositions 1, 2 and Theorem 1.

*Remark 3.* When the ordering  $f$  favors the existing tenants, these agents leave before every new applicant in the algorithm. Once they do, the algorithm reduces to a serial dictatorship. Therefore in such applications the mechanism can be used at two different points in time; first for the existing tenants and next for the new applicants.

## 6. EXTENSIONS

There are two assumptions in our model that seem to limit the practical significance of our results:

1. The null house  $h_0$  is assumed to be the last choice for all agents. While this assumption is natural for some applications (such as office space allocation for faculty), it may not be so natural in others. For example, a student may prefer off-campus housing to some on-campus houses. Moreover, some students may not even be eligible for some on-campus houses.

2. In many real-life applications there are several types of housing and agents have preferences over types of houses rather than houses themselves. By assuming strict preferences, we rule out such applications.

Nevertheless, these assumptions are made merely to simplify the notation and the exposition. Both assumptions can be dropped and our proposed mechanisms can be naturally extended to accommodate these more general models. We do this in Sections 6.1 and 6.2.

Another generalization that deserves attention is one that allows for consumption externalities. We discuss some difficulties associated with such a model in Section 6.3.

### 6.1. *What If the Null House Is Not the Last Choice?*

Suppose the null house  $h_0$  is not necessarily the last choice. In this case, for any given ordering  $f$ , we can modify the top trading cycles algorithm as follows: Introduce a *personalized null house*  $h_{0i}$  for each agent  $i$ . (It represents the no-house option for agent  $i$ ). In the modified algorithm, each personalized null house  $h_{0i}$  points to agent  $i$  until he or she leaves the market. At this step the null house  $h_{0i}$  leaves the market as well. Everything else stays the same in the algorithm. Note that if at any step a personalized null house  $h_{0i}$  is in a cycle, this cycle should be  $(i, h_{0i})$ . In this case, agent  $i$  prefers the null house to every remaining house and leaves the market without being assigned to a real house.

### 6.2. *What If There Are Multiple Units of a Housing Type?*

Suppose there are several types of housing and agents have preferences over types of housing rather than over houses. We assume that agents are indifferent between houses of the same type and they have strict preferences over types of houses. Since we need strict preferences in the top trading cycles algorithm, we will use a tie-breaking rule to eliminate these indifferences. Fortunately there is a natural tie-breaking rule in this context. Given any such preference profile  $R$  and ordering  $f$  of agents, construct a strict preference profile  $P$  as follows: For any agent  $i$ ,

1. given two houses of different types, the house of the better type (under  $R_i$ ) is strictly preferred under  $P_i$ ,
2. given two houses of the same type,
  - (a) if both houses are occupied then the house with the more senior owner (under  $f$ ) is strictly preferred under  $P_i$ ,
  - (b) if one is occupied and the other is vacant, then the occupied house is strictly preferred under  $P_i$ , and
  - (c) if both houses are vacant then the house with the lower index is preferred.

(This last case may look arbitrary but it is inessential since all vacant houses point to the same agent in the top trading cycles algorithm.)

Once the tie-breaking is handled the top trading cycles mechanism can be used.

### 6.3. What About Allowing for Consumption Externalities?

Suppose that agents not only care for their houses, but also for their neighbors, where their friends live, etc. Then one needs to allow for consumption externalities. Without adding additional structure in the model, all that can be done is allowing for any preference relation over the set of matchings. But such a model is formally equivalent to abstract models analyzed by Gibbard [6, 7] and Satterthwaite [19]. Therefore the only matching rules that are strategy-proof and Pareto efficient are *dictatorial* rules (Gibbard [6] and Satterthwaite [19]). Moreover the only lottery rules that are strategy-proof and ex post Pareto efficient are convex combinations of dictatorial rules (Gibbard [7]).

Another difficulty in a model with consumption externalities is formulating the individual rationality condition: While existing tenants have prior claims for their houses, this itself does not determine their welfare level in a model with consumption externalities.

## A. APPENDIX: PROOF OF THEOREM 2

Let  $\varphi$  be a mechanism that is Pareto efficient, individually rational, and strategy-proof. We want to show that, for all  $t \in \{1, \dots, |I|\}$ ,

$$\begin{aligned} & [\exists P \in \mathcal{P} \text{ s.t. } \varphi_{f(t)}(P) P_{f(t)} \varphi_{f(t)}^f(P)] \\ & \Rightarrow [\exists \tilde{P} \in \mathcal{P} \text{ and } s < t \text{ s.t. } \varphi_{f(s)}^f(\tilde{P}) \tilde{P}_{f(s)} \varphi_{f(s)}(\tilde{P})] \end{aligned} \quad (1)$$

We proceed by induction. Label agents so that  $i_t = f(t)$  for all  $t \in \{1, \dots, |I|\}$ .

*Claim 1.*  $\forall P \in \mathcal{P}, \varphi_{i_1}^f(P) R_{i_1} \varphi_{i_1}(P)$ .



Consider any ordering  $\sigma \in \Sigma$ . Recall that  $i_1 \notin J$  by construction,  $h^* \in G$  by relation (4), and  $h^* = \varphi_{i_1}(P) = \varphi_{i_1}(P^{|\mathcal{J}|}(\sigma))$  by construction. Moreover  $\bigcup_{j \in J} \varphi_j(P^0(\sigma)) = G$  by relation (6). Therefore, while all houses in  $G$  are assigned to agents in  $J$  by the mechanism  $\varphi$  under  $P^0(\sigma)$ , the same is not true under  $P^{|\mathcal{J}|}(\sigma) = P$ . Let  $P^{n(\sigma)}(\sigma)$  be the first member of the sequence  $P^0(\sigma), P^1(\sigma), \dots, P^{|\mathcal{J}|}(\sigma)$  under which an agent in  $J$  is assigned a house that is not in  $G$  by the mechanism  $\varphi$ . Note that  $1 \leq n(\sigma) \leq |\mathcal{J}|$  for all  $\sigma \in \Sigma$ .

We do not choose an arbitrary ordering. Pick an ordering  $\sigma^* \in \Sigma$  with

$$n(\sigma^*) = \min_{\sigma \in \Sigma} n(\sigma)$$

Note that such an ordering does not need to be unique. By the definition of  $n(\sigma^*)$ ,

$$\exists k \in J \quad \text{and} \quad h \in H \setminus G \quad \text{such that} \quad \varphi_k(P^{n(\sigma^*)}(\sigma^*)) = h. \quad (8)$$

Moreover relation (5), together with the individual rationality of  $\varphi$ , ensures that  $P_k^{n(\sigma^*)}(\sigma^*) = P_k$ . That is, agent  $k \in J$  should be revealing the preference relation  $P_k$  under the profile  $P^{n(\sigma^*)}(\sigma^*)$ . Now suppose agent  $k$  announces the preference relation  $P'_k$  instead. That is, consider the preference profile  $(P_{-k}^{n(\sigma^*)}(\sigma^*), P'_k)$ . Note that the number of agents in  $J$  revealing the preference relation  $P_j$  (rather than  $P'_j$ ) is  $n(\sigma^*) - 1$  under this profile. Therefore minimality of  $n(\sigma^*)$  ensures

$$\bigcup_{j \in J} \varphi_j(P_{-k}^{n(\sigma^*)}(\sigma^*), P'_k) = G$$

and this together with Pareto efficiency of  $\varphi$  ensures

$$\forall j \in J, \quad \varphi_j(P_{-k}^{n(\sigma^*)}(\sigma^*), P'_k) = \mu(j).$$

In particular, this is true for agent  $k \in J$ :

$$\varphi_k(P_{-k}^{n(\sigma^*)}(\sigma^*), P'_k) = \mu(k). \quad (9)$$

Relations (3), (8), and (9) imply

$$\varphi_k(P_{-k}^{n(\sigma^*)}(\sigma^*), P'_k) P_k \varphi_k(\underbrace{P_{-k}^{n(\sigma^*)}(\sigma^*), P_k}_{P^{n(\sigma^*)}(\sigma^*)}),$$

contradicting the strategy-proofness of  $\varphi$ . This completes the proof of Claim 1.



*Claim 2.* Let  $\ell \leq |I|$  and suppose relation (1) holds for all  $t \in \{1, \dots, \ell - 1\}$ . Then relation (1) holds for  $t = \ell$  as well.

*Proof of Claim 2.* Suppose relation (1) holds for all  $t \in \{1, \dots, \ell - 1\}$  but not for  $t = \ell$ . Then there exists a preference profile  $P \in \mathcal{P}$  such that  $\sigma_{i_\ell}(P) P_{i_\ell} \varphi_{i_\ell}^f(P)$  and

$$\forall \tilde{P} \in \mathcal{P}, \quad \forall t \in \{1, \dots, \ell - 1\}, \quad \varphi_{i_t}(\tilde{P}) \tilde{R}_{i_t} \varphi_{i_t}^f(\tilde{P}). \quad (10)$$

If there exists  $\tilde{P} \in \mathcal{P}$  and  $t \in \{1, \dots, \ell - 1\}$  such that  $\varphi_{i_t}(\tilde{P}) \tilde{P}_{i_t} \varphi_{i_t}^f(\tilde{P})$ , then by the induction hypothesis there exists  $P^* \in \mathcal{P}$  and  $s < t \leq \ell - 1$  such that  $\varphi_{i_s}^f(P^*) P_{i_s}^* \varphi_{i_s}(P^*)$  contradicting relation (10). Therefore

$$\forall \tilde{P} \in \mathcal{P}, \quad \forall t \in \{1, \dots, \ell - 1\}, \quad \varphi_{i_t}(\tilde{P}) = \varphi_{i_t}^f(\tilde{P}). \quad (11)$$

Denote  $\mu = \varphi^f(P)$  and  $h^* = \varphi_{i_\ell}(P)$ . Let  $J$  be the set of agents who leave the top trading cycles algorithm before agent  $i_\ell$  under  $P$ . Similarly let  $G$  be the set of houses that leave the algorithm before agent  $i_\ell$ . Define  $J_1 = J \setminus \{i_1, \dots, i_{\ell-1}\}$ . Note that no agent in  $I_N$ , unless he or she is a member of  $\{i_1, \dots, i_{\ell-1}\}$ , can leave the market before agent  $i_\ell$ . Therefore  $J_1 \subseteq I_E$ . We also have

$$\forall j \in J_1, \quad h_j \in G, \quad (12)$$

for otherwise agent  $j$  could not form a cycle before agent  $i_\ell$  left the market. Moreover,

$$\forall j \in J, \quad \forall h \in H \setminus G, \quad \mu(j) P_j h, \quad (13)$$

for otherwise agent  $j$  would not leave the market before house  $h$  did. Finally note that it is only houses in  $G$  that can be better for agent  $i_\ell$  than his or her assignment  $\mu(i_\ell)$ . Therefore

$$h^* \in G. \quad (14)$$

For all  $j \in J_1$ , let  $P'_j \in \mathcal{P}$  be such that

$$\forall h \in H \setminus \{\mu(j), h_j\}, \quad \mu(j) R'_j h_j P'_j h. \quad (15)$$

Observe that  $\varphi^f(P_{N \setminus J_1}, P'_{J_1}) = \varphi^f(P) = \mu$ . Therefore by relation (11) we have

$$\forall i \in \{i_1, \dots, i_{\ell-1}\}, \quad \varphi_i(P_{N \setminus J_1}, P'_{J_1}) = \varphi_i^f(P_{N \setminus J_1}, P'_{J_1}) = \varphi_i^f(P) = \mu(i)$$



By the definition of  $n(\pi^*)$ ,

$$\exists k \in J \quad \text{and} \quad h \in H \setminus G \quad \text{such that} \quad \varphi_k(P^{n(\pi^*)}(\pi^*)) = h. \quad (21)$$

Indeed, relation (20) implies that  $k \in J_1$ . Moreover, relation (13) together with the individual rationality of  $\varphi$  ensures that  $P_k^{n(\pi^*)}(\pi^*) = P_k$ . That is, agent  $k \in J_1$  should be revealing the preference relation  $P_k$  under the profile  $P^{n(\pi^*)}(\pi^*)$ .

Now suppose agent  $k$  announces the preference relation  $P'_k$  instead. That is, consider the preference profile  $(P^{n_{-k}(\pi^*)}(\pi^*), P'_k)$ . The number of agents in  $J_1$  revealing the preference relation  $P_j$  (rather than  $P'_j$ ) is  $n(\pi^*) - 1$  under this profile. Therefore minimality of  $n(\pi^*)$  ensures that

$$\bigcup_{j \in J} \varphi_j(P^{n_{-k}(\pi^*)}(\pi^*), P'_k) = G$$

and this together with relation (20) and Pareto efficiency of  $\varphi$  ensures that

$$\forall j \in J, \quad \varphi_j(P^{n_{-k}(\pi^*)}(\pi^*), P'_k) = \mu(j).$$

In particular, this is true for agent  $k \in J$ :

$$\varphi_k(P^{n_{-k}(\pi^*)}(\pi^*), P'_k) = \mu(k). \quad (22)$$

Relations (13), (21), and (22) imply

$$\varphi_k(P^{n_{-k}(\pi^*)}(\pi^*), P'_k) P_k \underbrace{\varphi_k(P^{n_{-k}(\pi^*)}(\pi^*), P_k)}_{P^{n(\pi^*)}(\pi^*)},$$

contradicting strategy-proofness of  $\varphi$  and completing the proof of Claim 2.

Claims 1 and 2, complete the proof. ■

### REFERENCES

1. A. Abdulkadiroğlu and T. Sönmez, Random serial dictatorship and the core from random endowments in house allocation problems, *Econometrica* **66** (1998), 689–701.
2. A. Bogomolnaia and H. Moulin, A simple random assignment problem with a unique solution, mimeo, University of Nottingham and Duke University, 1999.
3. S. Collins and K. Krishna, “The Harvard Housing Lottery: Rationality and Reform,” working paper, Georgetown University and Penn State University, 1995.
4. D. Gale and L. Shapley, College admissions and the stability of marriage, *Amer. Math. Monthly* **69** (1962), 9–15.
5. R. Garratt and C-Z. Qin, Cores and competitive equilibria with indivisibilities and lotteries, *J. Econ. Theory* **68** (1996), 531–543.

6. A. Gibbard, Manipulations of voting schemes: A general result, *Econometrica* **41** (1973), 587–601.
7. A. Gibbard, Manipulation of schemes that mix voting with chance, *Econometrica* **45** (1977), 665–681.
8. A. Hylland and R. Zeckhauser, The efficient allocation of individuals to positions, *J. Polit. Econ.* **87** (1979), 293–314.
9. D. E. Knuth, An exact analysis of stable allocation, *J. Algorithms* **20** (1996), 431–442.
10. V. Krishna and M. Perry, “Efficient Mechanism Design,” working paper, Penn State University, 1997.
11. J. Ma, Strategy-proofness and the strict core in a market with indivisibilities, *Int. J. Game Theory* **23** (1994), 75–83.
12. S. Mongell and A. E. Roth, Sorority rush as a two-sided matching mechanism, *Amer. Econ. Rev.* **81** (1991), 415–440.
13. S. Papai, Strategyproof assignment by hierarchical exchange, mimeo, Koç University, 1998.
14. A. E. Roth, Incentive compatibility in a market with indivisibilities, *Econ. Lett.* **9** (1982), 127–132.
15. A. E. Roth, The evolution of labor market for medical interns and residents: A case study in game theory, *J. Polit. Econ.* **92** (1984), 991–1016.
16. A. E. Roth, A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in U.K., *Amer. Econ. Rev.* **81** (1991), 441–464.
17. A. E. Roth and E. Peranson, The effects of a change in the NRMP matching algorithm, *J. Amer. Medical Assoc.* **278** (1997), 729–732.
18. A. E. Roth and A. Postlewaite, Weak versus strong domination in a market with indivisible goods, *J. Math. Econ.* **4** (1977), 131–137.
19. M. A. Satterthwaite, Strategy-proofness and Arrow’s conditions: Existence and correspondence theorems for voting procedures and social welfare functions, *J. Econ. Theory* **10** (1975), 187–216.
20. L. Shapley and H. Scarf, On cores and indivisibility, *J. Math. Econ.* **1** (1974), 23–28.
21. L. G. Svensson, Strategy-proof allocation of indivisible goods, mimeo, Lund University, 1998. [Forthcoming in *Social Choice and Welfare*]
22. W. Thomson, Consistency in exchange economies, mimeo, University of Rochester, 1992.
23. L. Zhou, On a conjecture by Gale about one-sided matching problems, *J. Econ. Theory* **52** (1991), 123–135.