



# Designing mechanisms to focalize welfare-improving strategies <sup>☆</sup>



Daniel E. Fragiadakis <sup>a,\*</sup>, Peter Troyan <sup>b</sup>

<sup>a</sup> Villanova University, United States of America

<sup>b</sup> University of Virginia, United States of America

## ARTICLE INFO

### Article history:

Received 9 February 2017

Available online 12 February 2019

### JEL classification:

C78

C91

C92

D61

D63

### Keywords:

Dictatorship

Indifference

Matching

Welfare

Efficiency

Experiments

## ABSTRACT

Many institutions use matching algorithms to allocate resources to individuals. Examples include the assignment of doctors, students and military cadets to hospitals, schools and branches, respectively. Oftentimes, agents' ordinal preferences are highly correlated, motivating the use of mechanisms that provide agents with channels through which they can express some cardinal preference information. This paper studies two such mechanisms, one from the field and one we design. In each of the games induced by these algorithms, we identify the strategies that constitute the unique symmetric ex-post equilibrium. Interestingly, when we test the mechanisms in the lab, these equilibrium predictions fail. Subjects nevertheless behave largely in concordance with the mechanisms' *intended* strategies; the focalization of such strategies lead to greater welfare in relation to a popular existing mechanism.

© 2019 Elsevier Inc. All rights reserved.

## 1. Introduction

The *assignment problem* (also sometimes called the *house allocation problem*, following the nomenclature of the seminal paper of Shapley and Scarf, 1974) is a standard model of allocation of indivisible resources to agents without the use of monetary transfers. Real-world examples include assigning students to seats in public schools, public housing units to tenants, cadets to military branches, organ donations to recipients and undergraduates to university housing.<sup>1</sup> Less formal examples include assigning workers to projects, professors to offices or children to household chores.

<sup>☆</sup> We would like to thank Sandro Ambuehl, Blake Barton, Douglas Bernheim, Itay Fainmesser, Clayton Featherstone, Simon Gächter, Guillaume Haeringer, Jeffrey Naecker, Krishna Rao and Charles Sprenger for helpful suggestions, as well as Liam Anderson and Kimberley Williams from the Stanford Graduate School of Business. Detailed comments from an editor and two anonymous reviewers that have substantially improved the paper are also greatly appreciated. We are especially grateful to Fuhito Kojima, Muriel Niederle and Al Roth for many conversations related to this project. Generous support from the Leonard W. Ely and Shirley R. Ely Fellowship (Fragiadakis) and B.F. Haley and E.S. Shaw Fellowship (Troyan), both through the Stanford Institute for Economic Policy Research, is gratefully acknowledged.

\* Corresponding author.

E-mail addresses: [danielfragiadakis@gmail.com](mailto:danielfragiadakis@gmail.com) (D.E. Fragiadakis), [troyan@virginia.edu](mailto:troyan@virginia.edu) (P. Troyan).

<sup>1</sup> See, for example, Thakral (2015), Sönmez and Switzer (2013), Sönmez (2013), Roth et al. (2004), Chen and Sönmez (2002; 2004), Sönmez and Ünver (2005 and 2010b).

In many assignment problems, agents' *ordinal* preferences are highly correlated. For instance, in school choice markets, many high-quality schools are over-demanded (see Abdulkadiroğlu et al., 2011 and 2015). In such competitive settings, eliciting preference information from agents that is purely ordinal can leave efficiency gains on the table. For example, consider the following simple assignment problem of matching three goods to three agents in a one-to-one fashion (which also appears in Abdulkadiroğlu et al., 2011); the numbers in the chart represent agents' cardinal utilities for the goods.

	good 1	good 2	good 3
agent 1	0.8	0.2	0
agent 2	0.6	0.4	0
agent 3	0.8	0.2	0

Suppose the popular Random Serial Dictatorship (RSD) mechanism is used to solve the assignment problem. In RSD, agents submit strict rankings over the goods and are sequentially assigned to their favorite available good according to a randomly chosen *priority*. It can be readily shown that RSD has a Dominant-Strategy Equilibrium (DSE) where each agent states her true rankings.<sup>2</sup> In this example, the DSE induces all agents to submit the same ordinal ranking, resulting in expected utilities of 1/3 for each agent. A more efficient allocation (in an *ex ante* sense) would be to give good 2 to agent 2 with probability 1 and give agents 1 and 3 each a 50/50 chance of obtaining goods 1 and 3, resulting in expected utility of 0.4 for each agent.<sup>3</sup>

The ex-ante efficiency limitations of RSD arise because agents can express only their *ordinal* preferences. The goal of this paper is to investigate the extent to which it is possible to raise welfare by designing a mechanism that allows agents to communicate *cardinal* preference information using an ordinal reporting language.<sup>4</sup> Indeed, this issue arises in practical applications such as school choice (Abdulkadiroğlu et al., 2011 and 2015) and teacher assignment (Coffman et al., 2017 and Featherstone, 2014). Such an approach is also implemented at a world-class MBA program to assign its students to educational trips. The mechanism we design in this paper is inspired by this program's algorithm as well as RSD.

We refer to the MBA program's mechanism as "Popularity" because it prioritizes students who highly rank trips that, according to the set of all students' submitted preferences, are ranked lowly on average, i.e., are unpopular. Popularity also allows agents to submit *weak* ordinal rankings. The mechanism's designers intuited that this would induce participants to follow their cardinal utilities when stating their preferences. Specifically, they thought a higher-valued good would be ranked weakly better than a lower-valued one and if two goods were sufficiently close in value, they would be placed into the same indifference class, or "bin" for short. We call such behavior *Cardinal Following* (CF).

The reason to expect that Popularity might focalize CF reporting is that a CF strategy balances the tradeoffs of stating an indifference over two similarly valued goods. The benefit of doing so is that one can receive a weakly better priority. The cost is that it gives the mechanism more leeway when choosing what good to ultimately assign to an agent from within a particular bin. For example, in the aforementioned market, suppose agents all placed goods 1 and 3 in their first and second bins, respectively, but follow their cardinal utilities when deciding where to place good 2. Agents 1 and 3, whose values for good 2 are relatively low, place good 2 in their second bins, making the report " $\{1\} > \{2, 3\}$ ", while agent 2, whose value for good 2 is relatively high, places good 2 into her first bin, making the report " $\{1, 2\} > \{3\}$ ". Under this profile of submitted reports, Popularity gives agent 2 the highest priority (for highly ranking the unpopular good 2) and guarantees agent 2 *some* good from her first bin, i.e., good 1 or 2. Agent 2 then ultimately obtains good 2 once Popularity allocates good 1 to the agent who is prioritized second. Each agent then attains expected utility equal to 0.4 (versus 1/3 in the DSE of RSD).

Note, however, that in this example, the given strategy profile is actually *not* an equilibrium: agent 1 can deviate to reporting  $\{1, 2\} > 3$  and be made better off. In a more general model, we solve for Popularity's *unique* symmetric Ex-Post Equilibrium (EPE), which is to place all but the least-preferred good into one's first indifference class, irrespective of cardinal preferences. We show that the EPE yields welfare that is no better than that in the DSE of RSD. Whether real-world participants are likely to follow the predicted equilibrium strategies, however, is ultimately an empirical question.<sup>5</sup> Perhaps such individuals will behave in ways that are more consistent with the Cardinal Following strategies that Popularity seeks to elicit. Since actual play is what ultimately determines the level of welfare achieved in practice, a thorough empirical investigation of Popularity is necessary. Using existing field data presents difficulties because we do not observe the participants' true utilities; thus, we turn to a lab experiment.

In our *Popularity* treatment (that assigns goods to subjects using Popularity), 34% of reports are CF and less than 1% are the EPE (despite subjects receiving feedback between decisions). Earnings, our welfare measure, are greater on average

<sup>2</sup> It is well-known that RSD is also ex-post Pareto efficient and fair, further rationalizing its common use. Pycia and Troyan (2016) provide the first formal characterization of RSD as the *unique* mechanism that is ex-post efficient, fair, and obviously strategy-proof (Li, 2017). See also Abdulkadiroğlu and Sönmez (1998), Bogomolnaia and Moulin (2001), and Liu and Pycia (2016) for other formal discussions of RSD.

<sup>3</sup> These gains from trade stem from the competition for the goods. In practice, we often observe correlation among agents' ordinal preferences.

<sup>4</sup> While it is possible to attempt to elicit cardinal preferences directly, it is much easier for real-world agents to report ordinal preference information, and in practice, most mechanisms use ordinal reporting languages. This is discussed further in Section 5.

<sup>5</sup> A plethora of lab evidence documents departures from equilibrium predictions (see Camerer, 2003 for an overview).

compared to the DSE of RSD; however, at an individual level, 29% of subjects actually fare worse under Popularity, with some faring significantly worse. We find that these subjects' reports often contain *Preference Reversals* (PRs), i.e., involve stating a good as being strictly preferred to another despite the latter being of strictly greater value.<sup>6</sup> Thus, while Popularity raises welfare on average, one of the two main types of behavior it focalizes actually tends to backfire for the agents who follow it and makes them worse off.

In search of more uniform welfare improvements, we attempt to modify Popularity in a way that shifts behavior away from PRs and towards CF reporting. In particular, we design an "Indifference" mechanism that prioritizes an agent exclusively for stating indifference, irrespective of how highly she ranks unpopular goods. We find that Indifference is much more effective in focalizing CF reporting in comparison to Popularity: CF behavior increases from 34% to 61%. In addition, Indifference causes the number of reports containing PRs to drop substantially in relation to Popularity: the percentages of such reports goes from 28% down to 3%. The EPE is still played less than 1% of the time under Indifference.

More important than behavior, however, is the ultimate level of welfare generated by Indifference. We find that it yields significantly greater welfare on average compared to the DSE of RSD. Compared to Popularity, Indifference yields similar aggregate welfare. A case can be made for using Indifference instead of Popularity, however, if the *distribution* of welfare is considered. Compared to the DSE of RSD, 95% of participants earn more under Indifference; under Popularity, this number is only 71%, with some participants earning significantly less than they would in the DSE of RSD.

The final question we consider in this paper is how much Indifference's welfare improvements over RSD come from its incentives versus simply the *opportunity* for stating weak preferences. Even with unique valuations for goods and no explicit incentives for stating weak preferences, individuals may express some indifference out of altruism, for example, or environmental cues, in which case a strategy-proof "Random" mechanism that prioritizes agents at random but is otherwise identical to Indifference may be ideal. In our experiment, a subject always values each good distinctly, and thus has a dominant strategy of stating no indifference if her goal is to maximize her own expected earnings. We find, however, this behavior describes only 21% of reports in the *Random* treatment. Since ordinal preferences are common, the remaining 79% of non-truthful reports give off positive externalities, enough, in fact, to bring average earnings beyond their equilibrium level. Subjects still earn less, however, than what they earn under Indifference. Thus, allowing weak rankings raises welfare slightly, but incentivizing them brings even greater gains.

In sum, this paper follows the spirit of Roth (2002), who argues that market design should be treated as a form of "economic engineering" in which theoretical insights must be combined with additional computational, empirical, and experimental tools for successful practical implementation. More and more, we see economic literature and policy that takes agents' limited rationality under consideration; as market designers, it is important that we too acknowledge the cognitive limitations of real-world agents and design mechanisms that are as user-friendly as possible. We do so via Indifference by focalizing non-truthful, yet welfare-improving behavior. Our approach shows that the theoretical gold-standard of *strategy-proofness* (dominant strategy incentive compatibility) may not always be necessary for predicting behavior and may unnecessarily limit efficiency in practice.

At the same time, not all manipulable (non-strategy-proof) mechanisms are created equal, as seen by the differences in results from Indifference and Popularity; other work has shown that some popular manipulable mechanisms may not induce the desired welfare-improving behavior in practice (see Section 5). The main takeaway of our results is that mechanisms that are *designed* to be manipulable can raise welfare in practice, but only if they are carefully crafted and tested to focalize the appropriate strategies. We show one way that this can be done in a canonical setting that has numerous potential applications. More generally, we believe that this type of "behavioral market design" is a fruitful area for future work.

The remainder of this paper is organized as follows. Section 2 presents the model, mechanisms and the EPE. Section 3 outlines the experiment whose results we report in Section 4. Section 5 includes related literature and a discussion.

## 2. Model

There is a set of goods  $A = \{1, \dots, n\}$  and a set of individuals  $N = \{i_1, \dots, i_n\}$ , each of whom demands exactly one good. A *mechanism* is a function that takes as an input a collection of messages from the agents and outputs an *allocation*, which is simply an assignment that gives each agent exactly one of the goods. A message is an ordered partition of  $A$ . We call each member of the partition a "bin" and let  $B_i^j$  denote the set of goods that  $i$  places in bin  $j$ . Thus, to represent the submitted message  $x_i$  of agent  $i$ , we write  $x_i = \{B_i^1 \succ_i \dots \succ_i B_i^n\}$ , with the interpretation that  $B_i^1$  is agent  $i$ 's highest ranked bin,  $B_i^2$  is  $i$ 's second highest ranked bin, and so forth. To ensure  $x_i$  is a partition of  $A$ , we require  $\bigcup_{k=1}^n B_i^k = A$  and  $B_i^y \cap B_i^z = \emptyset$  for all  $y \neq z$  (note that some bins  $B_i^k$  may be empty). Formally, we impose one additional restriction on reports:  $B_i^y \neq \emptyset \implies B_i^{y-1} \neq \emptyset$  for all  $y > 1$ . This says that bins cannot be "skipped".

<sup>6</sup> While PRs are not dominated strategies, the problem is that while highly ranking an unpopular good  $g$  gives a subject high priority, she is often ultimately assigned  $g$ . This causes subjects who submit reports that contain PRs to receive (much) lower earnings on average. We find that 28% of submitted reports contain PRs, and, further, subjects do not seem to learn from their mistakes: 38% of participants state reports containing PRs more than half the time despite the receipt of feedback.

2.1. The mechanisms

The main mechanisms we study in this paper are Popularity and Indifference.<sup>7</sup> Both use agent priorities as a first step to making assignments. The priority rules of Popularity and Indifference are given in the boxed descriptions below.

**Generating  $\succeq_P$ , the weak ordering of agents under Popularity:**

**Step 1)** For each good, count the number of agents who place it in their *first bins*. In other words, compute each good  $g$ 's demand as  $|\{i : g \in B_i^1 \text{ for some agent } i\}|$ .

**Step 2)** Order the goods strictly from most to least demanded, breaking ties randomly. Call the resulting ordering good *desirability*. (See Figs. 1 and 2 for an example.)

**Step 3)** For each pair of agents,  $i$  and  $j$ , find the smallest integer  $k$  such that  $B_i^k \neq B_j^k$ . If such a  $k$  exists,  $g \in B_i^k \implies i \succeq_P j$  and  $g \in B_j^k \implies j \succeq_P i$ , where  $g$  is the *least desirable* good in  $(B_i^k \cup B_j^k) - (B_i^k \cap B_j^k)$ . If no such  $k$  exists, then  $i \succeq_P j$  and  $j \succeq_P i$ .

**Generating  $\succeq_I$ , the weak ordering of agents under Indifference:**

For each pair of agents,  $i$  and  $j$ , find the smallest integer  $k$  such that  $|B_i^k| \neq |B_j^k|$ . If such a  $k$  exists,  $|B_i^k| > |B_j^k| \implies i \succeq_I j$  and  $|B_j^k| > |B_i^k| \implies j \succeq_I i$ . If no such  $k$  exists, then  $i \succeq_I j$  and  $j \succeq_I i$ .

Once  $\succeq_P$  and  $\succeq_I$  are established, each is used to generate a strict weak ordering of agents, breaking ties uniformly at random. Each algorithm then follows a procedure of *designating* a bin to each agent. A subset  $N' \subseteq N$  of individuals' bin designations are *feasible* if and only if there exists an allocation such that each individual in  $N'$  receives *some* good from her designated bin. The bin designation rules for Popularity and Indifference are given in the boxed descriptions below.

**Bin Designation Rule of Popularity:**

- 1) Designate bin  $B_1^1$  to agent  $i_1$ . (This is always feasible.) Let  $i_j$  denote the agent most recently designated a bin.
- 2) If  $i_j$  is designated  $B_j^1$ , find the smallest  $j' > j$  such that  $i_{j'}$  can be feasibly designated  $B_{j'}^1$ , and designate  $B_{j'}^1$  to  $i_{j'}$ . If no such  $j'$  exists, find the highest priority agent without a bin designation and designate her most preferred bin to her, subject to feasibility.
- 3) If  $i_j$  is designated  $B_j^k$  for some  $k > 1$ , find the smallest  $j' > j$  such that  $i_{j'}$  is without a bin designation and designate her most preferred bin to her, subject to feasibility. Once no such  $j'$  exists, all agents have been designated bins.

**Bin Designation Rule of Indifference:**

- 1) Designate bin  $B_1^1$  to agent  $i_1$ . (This is always feasible.) Let  $i_j$  denote the agent most recently designated a bin. If  $j = n$ , all agents have been designated bins.
- 2) If  $j < n$ , designate agent  $i_{j+1}$ 's most preferred bin to her, subject to feasibility.

Upon completion of the bin designation process, each agent is allocated a good from her designated bin.<sup>8</sup> Note that final allocations are ex-post efficient with respect to the bin designations: if a reallocation gives agent  $i_j$  a good from a bin more preferred than her designated bin, another agent  $i_{j'}$  must be given a good from a bin preferred less than her designated bin. Figs. 1 and 2 provide examples of Indifference and Popularity.<sup>9</sup>

<sup>7</sup> We compare these theoretically and empirically to the Random Serial Dictatorship (RSD). In an additional control treatment (which we call *Random*), we also test a mechanism that is equivalent to Indifference in all aspects, except that it orders agents at random. We refer to this mechanism as "Random".

<sup>8</sup> Under Popularity and Indifference, a final set of bin designations may generate multiple feasible allocations. Each such allocation is equally likely to be realized in each mechanism.

<sup>9</sup> In designing Indifference, we sought a mechanism that real-world agents would find as "user-friendly" as possible. To do so, we simplified Popularity both in terms of how (i) agents are prioritized and (ii) bins are designated. To fully understand the individual effects (i) and (ii) would require testing two additional mechanisms that modify Popularity by only changing (i) and (ii), respectively.

Indifference					
agent	$B_i^1$	$B_i^2$	$B_i^3$	$B_i^4$	$B_i^5$
$i = 1$	$g_1, g_2, G_4$	$g_3, g_5$			
$i = 2$	$G_1, g_2$	$g_3, g_5$	$g_4$		
$i = 3$	$G_2$	$g_3, g_4$	$g_1$	$g_5$	
$i = 4$	$g_2$	$g_1$	$G_3, g_4$	$g_5$	
$i = 5$	$g_3$	$g_1$	$g_2$	$g_4$	$G_5$

Popularity					
agent	$B_i^1$	$B_i^2$	$B_i^3$	$B_i^4$	$B_i^5$
$i = 1$	$g_1, g_2, G_4$	$g_3, g_5$			
$i = 5$		$G_3$	$g_1$	$g_2$	$g_4$
$i = 2$	$G_1, g_2$	$g_3, g_5$	$g_4$		
$i = 3$	$G_2$	$g_3, g_4$	$g_1$		$g_5$
$i = 4$	$g_2$	$g_1$	$g_3, g_4$		$G_5$

**Fig. 1.** The left and right panels – corresponding to Indifference and Popularity, respectively – show the agent orderings and assignments made from a single set of reports. Each “agent” column signifies the agent ordering (where higher agents are ordered earlier). Each cell indicates the composition of an agent’s bin. (The reports of the agents are the same in both tables, but the agent orderings and final assignments are different.) For instance, in each panel, the cell intersected by the  $i = 3$  row and  $B_i^2$  column indicates that agent 3 places goods 3 and 4 in her second bin. The good that an individual receives is indicated by a capital letter. For example, under both mechanisms, individuals 1, 2 and 3 receive goods 4, 1 and 2, respectively. Boxes and circles are drawn around individuals and goods that are matched differently across the two mechanisms. The agent ordering under Popularity is produced using desirability 1 from Fig. 2. (It can be verified that the assignments would not change under desirability 2 from Fig. 2.)

Good	$g_5$	$g_4$	$g_3$	$g_1$	$g_2$
Demand	0	1	1	2	4
Desirability 1	1	$\boxed{2}$	$\textcircled{3}$	4	5
Desirability 2	1	$\textcircled{3}$	$\boxed{2}$	4	5

**Fig. 2.** Using the reports from Fig. 1, we compute the demand for each good. For example, since there are 4 individuals who place good 2 in their first bins, the demand of good 2 is 4. Given that goods 3 and 4 have the same demands, there are two definitions of desirability that can be used to break ties (desirability 1 and desirability 2). Boxes and circles show the distinction between the two desirability definitions.

2.2. Ex-post equilibrium (EPE)

Each of the mechanisms described above induces a game amongst the agents. To analyze the equilibria of these games (and the resulting welfare), we must model the agents’ (cardinal) preferences. Let  $v_i(k)$  denote agent  $i$ ’s cardinal utility from good  $k$ . We assume that each agent draws a vector of cardinal utilities  $v_i = (v_i(1), \dots, v_i(n))$  from the set  $V_i = \{v_i : 1 = v_i(1) > v_i(2) > \dots > v_i(n) = 0\}$  according to some CDF  $F$ , independently across agents. In other words, agents have strict, common ordinal preferences, and the cardinal utilities of the best and worst goods are normalized to 1 and 0, respectively.<sup>10</sup> We will sometimes call the vector  $v_i$  agent  $i$ ’s type. Under any mechanism, a strategy for agent  $i$  is then a mapping  $s_i : V_i \rightarrow X_i$  from types to messages. We let  $s = (s_1, \dots, s_n)$  denote a strategy profile. A strategy profile  $s$  is symmetric if  $s_i = s_j$  for all  $i, j \in N$ .

The strongest equilibrium concept is equilibrium in dominant strategies. If a mechanism has a dominant strategy equilibrium, then each agent has a strategy  $s_i^*$  that is (weakly) better than all alternatives for any opponent play. Dominant strategy equilibria are generally very restrictive, and often do not exist. In our case, RSD and Random<sup>11</sup> will have the same dominant strategy equilibrium (for each  $v_i$ ,  $s_i^*(v_i)$  is the strict ordinal preference relation associated with  $v_i$ ), but Popularity and Indifference will not.

We thus instead focus on ex-post equilibrium. Let  $u_i(v_i, x)$  be  $i$ ’s expected utility when she is of type  $v_i$  and the agents submit the message profile  $x$  to the mechanism.<sup>12</sup> Formally, we have the following definition.

**Definition.** A profile of strategies  $s = (s_1, \dots, s_n)$  is an **ex-post equilibrium** if for all  $i$ ,  $v_i \in V_i$ ,  $v_{-i} \in V_{-i}$  and  $x_i \in X_i$ , the following holds:

$$u_i(v_i, s_i(v_i), s_{-i}(v_{-i})) \geq u_i(v_i, x_i, s_{-i}(v_{-i})).$$

In words, an ex-post equilibrium is one in which no agent would want to change her strategy, even if she knew the private information (types) of her opponents. Ex-post equilibrium as a solution concept has received a great deal of attention recently in the mechanism design literature in settings where dominant strategy equilibria either may not exist or require

<sup>10</sup> In the lab experiment to follow, giving subjects common ordinal – but different cardinal – preferences gives us the cleanest possible test of how successful our mechanisms are at communicating cardinal utility using an ordinal reporting language, and so we make this assumption here as well. Additionally, environments with high ordinal preference correlation are common empirically (see, e.g., Abdulkadiroğlu et al. (2011), who also make a common ordinal preference assumption in a school choice environment), and it is in these settings where our mechanisms have the most potential benefit. Before using Popularity, the MBA program used RSD for only one year. RSD gave over 10% of students trips they ranked 8th or worse, suggesting high ordinal preference correlation.

<sup>11</sup> Random is the mechanism we test in our control treatment; it is the same as Indifference in all aspects except that it orders agents uniformly at random (and thus independently of their submitted reports).

<sup>12</sup> Note that the message profile is fixed at  $x$ , and the randomness comes from any potential tie-breaking when running the mechanism.

reporting a degree of information that is infeasible from a communication/computational perspective.<sup>13</sup> It is a weaker solution concept than dominant strategy equilibrium, but stronger than Bayesian equilibrium. In particular, ex-post equilibria are more robust to informational assumptions about the environment: in an ex-post equilibrium, an agent's strategy does not depend on her beliefs about others' types. This suggests that, when such equilibria exist, they should be more compelling as a prediction than Bayesian equilibria.<sup>14</sup>

We will show (in Theorems 1 and 2) that the Popularity and Indifference mechanisms both have a unique Ex-Post Equilibrium (EPE). Recall that the message space  $X_i$  for each agent  $i$  is the space of ordered partitions, or “bins”. Consider the following message,  $x_i^{AbL}$ , which we refer to as the “All but Last” message:  $x_i^{AbL}$  places goods 1, 2, ...,  $n - 1$  (all but the “last” good) into bin 1, and places good  $n$  into bin 2. Let  $s_i^{AbL}$  be the strategy profile such that  $s_i^{AbL}(v_i) = x_i^{AbL}$  for all  $v_i$ , i.e.,  $i$  always reports the AbL message.

**Theorem 1.** *The games induced by Popularity and Indifference each have a symmetric Ex-Post Equilibrium (EPE), given by the strategy profile  $s^{AbL} = (s_1^{AbL}, \dots, s_n^{AbL})$ .*

In the EPE, all agents' strategies “unravel” to the trivial AbL strategy in which no cardinal information is conveyed. As a result, Popularity and Indifference first order agents at random in the EPE. An agent thus has a  $1/n$  chance of being ordered last and being assigned good  $n$ . With the remaining  $(n - 1)/n$  probability, the agent will be designated her first bin. Conditional on being designated her first bin, she has a  $1/(n - 1)$  chance of obtaining any given good inside it. Thus, her unconditional probability of receiving each good is  $1/n$ , just as in the Dominant-Strategy Equilibrium (DSE) of the Random Serial Dictatorship (RSD).

Finally, we prove uniqueness of the EPE in the following theorem.

**Theorem 2.** *In each of the games induced by Popularity and Indifference, the **unique** symmetric ex-post equilibrium is the EPE given in Theorem 1.*

Whether real-world agents will ultimately play the EPE in the games induced by Popularity and Indifference is an important empirical question. It is entirely possible that it will not be focal in practice, and that behavior will be more consistent with other kinds of strategies, such as the Cardinal Following strategies that these mechanisms seek to elicit. To determine what behavior – and more importantly resulting welfare – these mechanisms yield in practice, we turn to a lab experiment.

### 3. Experiment

Our lab experiment was run in ztree (Fischbacher, 2007) in Spring 2013 at a large American university using mostly undergraduate subjects who participated in exactly one of three treatments: *Popularity*, *Indifference* and *Random*. The *Random* treatment serves as a control to test whether any potential welfare gains in the *Indifference* treatment come simply from allowing agents to state indifference. In the *Random* treatment, a “Random” mechanism is used to make assignments, where Random is equivalent to Indifference in all aspects except that it creates an agent priority ordering uniformly at random, irrespective of submitted reports.

Each treatment had 24 subjects who were further partitioned into fixed and strategically independent groups of 6. In treatment  $X$ , mechanism  $X$  was used to make allocations in 20 Assignment Games. In an Assignment Game, a player has a Payoff Function indicating her Experimental Currency Unit (ECU)<sup>15</sup> value of each good. Subjects are told that all participants value good  $i$  strictly more than good  $i + 1$  for all  $i = 1, \dots, 5$ .<sup>16</sup> We endow subjects, independently and uniformly at random, with Payoff Functions (Fig. 3). Players simultaneously place goods into bins to provide weak ordinal rankings over goods.<sup>17</sup>

Before each Assignment Game, Payoff Functions were redrawn and subjects had 60 seconds to submit their reports.<sup>18</sup> After each game, subjects were given feedback (for about 30 seconds) of all previous periods: they saw *their own* Payoff Functions, decisions and assignments. At the end of the 20 periods, each subject received the value of the good she was assigned in a randomly chosen period. Participants earned \$5 for showing up, \$11.47 on average from Assignment Games, and \$5 for completing the study. The experiment lasted about 75 minutes.

<sup>13</sup> For instance, Edelman et al. (2007) study the generalized second-price auction (GSP) used by Google to sell keywords. GSP does not have a dominant strategy equilibrium, but does have an ex-post equilibrium. Holzman and Monderer (2004) argue for using ex-post equilibrium as a solution concept in combinatorial auctions where the dominant strategy of reporting a full valuation function in a VCG mechanism is prohibitively complex to carry out in practice.

<sup>14</sup> Bergemann and Morris (2005 and 2012) make this point convincingly, and then study implementation in ex-post equilibrium. Further, note that any ex-post equilibrium will also be a Bayesian equilibrium.

<sup>15</sup> Four ECU are equivalent to one US Dollar.

<sup>16</sup> They are also told that goods 1 and 6 are commonly valued at 99 ECU and 3 ECU, respectively.

<sup>17</sup> The instructions for each treatment are shown in Appendix A.4.

<sup>18</sup> Time limits were soft; the experimenter merely told subjects to hurry up if the limits were exceeded.

	$g_1$	$g_2$	$g_3$	$g_4$	$g_5$	$g_6$		$g_1$	$g_2$	$g_3$	$g_4$	$g_5$	$g_6$
PF <sub>1</sub>	99	97	95	93	13	3	PF <sub>6</sub>	99	97	17	7	5	3
PF <sub>2</sub>	99	97	95	15	13	3	PF <sub>7</sub>	99	19	17	15	13	3
PF <sub>3</sub>	99	97	95	15	5	3	PF <sub>8</sub>	99	19	17	15	5	3
PF <sub>4</sub>	99	97	17	15	13	3	PF <sub>9</sub>	99	19	17	7	5	3
PF <sub>5</sub>	99	97	17	15	5	3	PF <sub>10</sub>	99	19	9	7	5	3

**Fig. 3.** Cells are color-coded to illustrate three Cardinal Following (CF) strategies: CF1 places goods from black, gray and white cells into bins 1, 2 and 3, respectively, CF2 places goods from black cells into bin 1 and the rest into bin 2 and CF3 places goods from black and gray cells into bin 1 and the rest into bin 2.

**4. Results**

4.1. Analysis of welfare

We begin our earnings analysis on a “market” level. For a particular Assignment Game – which depicts a matching market of 6 subjects – we compute three forms of earnings:

- ① *earnings<sub>Empirical</sub>*: the sum of subjects’ expected earnings, computed using the algorithm in the given Assignment Game and the subjects’ submitted reports
- ② *earnings<sub>RSD</sub>*: the sum of subjects’ expected earnings, computed using the Random Serial Dictatorship, assuming subjects play the Dominant-Strategy Equilibrium (DSE)
- ③ *earnings<sub>Soc.Plan</sub>*: the highest total earnings that a Social Planner could reach over all possible assignments (given subjects’ drawn Payoff Functions in the Assignment Game)

①, ② and ③ let us compute an Assignment Game’s “Scaled Earnings”:

$$\text{Scaled Earnings} = \frac{\text{earnings}_{\text{Empirical}} - \text{earnings}_{\text{RSD}}}{\text{earnings}_{\text{Soc.Plan}} - \text{earnings}_{\text{RSD}}} \tag{1}$$

Scaled Earnings (SE) allow us to evaluate welfare against two benchmarks simultaneously. When  $SE = 0$ , welfare is equal to that of RSD; when  $SE = 1$ , welfare is equal to that of a Social Planner. In addition, when comparing welfare across two treatments, SE control for the possibility that subjects in one treatment, on average, may have drawn better Payoff Functions than subjects in the other.<sup>19</sup>

In our experiment, subjects receive feedback after each Assignment Game. Hence, we do *not* treat each one as a statistically independent observation. Nonetheless, because each 6-subject market was completely strategically isolated from all the others (even within the same treatment), we can average each market’s SE over the 20 games (shown in Table 1) to build a set of statistically independent observations that we use to obtain our two results regarding aggregate welfare (Results 1 and 2).

**Result 1.** Mean Scaled Earnings in the Popularity, Indifference and Random treatments are each significantly greater in comparison to the welfare from the DSE in RSD (two-tailed Mann–Whitney tests at 5% significance).

**Result 2.** Mean Scaled Earnings in the Popularity and Indifference treatments are not significantly different from one another, yet both are significantly greater than the scaled earnings in the Random treatment (two-tailed Mann–Whitney tests at 5% significance).

While Results 1 and 2 document average welfare improvements when moving away from RSD, such a move could be hasty if it is far from Pareto-improving. We thus turn to an individual earnings analysis and compute two additional measures:

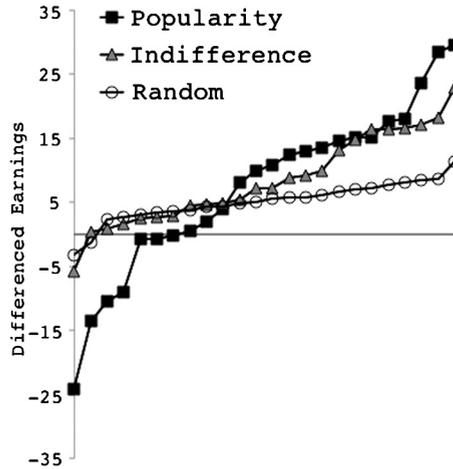
- ① *earnings<sub>Empirical</sub>*: subjects’ experimental earnings, averaged over all 20 games

<sup>19</sup> If one mechanism yields higher earnings than another, we do not want it to simply be because the former happened to endow subjects with better Payoff Functions due to randomness.

**Table 1**

The table shows each group's Scaled Earnings, averaged over the 20 games in the experiment. The final row in the table shows the mean of these averages.

Group	Treatments		
	Popularity	Indifference	Random
Highest	0.62	0.60	0.33
	0.49	0.57	0.32
	0.43	0.54	0.29
Lowest	0.34	0.35	0.27
Mean	0.47	0.51	0.30



**Fig. 4.** The figure shows each subject's Differenced Earnings which are nearly all positive in the *Indifference* and *Random* treatments, indicating that either of these would be an approximate Pareto improvement over RSD. Moving from RSD to *Popularity*, however, makes some subjects substantially better off (top right corner of graph), but at the expense of making some subjects substantially worse off (bottom left corner of graph). Lastly, the Differenced Earnings from the *Indifference* treatment roughly stochastically dominate those from the *Random* treatment, indicating that moving from *Random* to *Indifference* would be an approximate Pareto improvement.

②  $earnings_{RSD}$ : subjects' earnings under the Random Serial Dictatorship, assuming subjects play the Dominant-Strategy Equilibrium (DSE), averaged over all 20 games

① and ② let us compute an individual subject's "Differenced Earnings":

$$\text{Differenced Earnings} = earnings_{\text{Empirical}} - earnings_{RSD} \tag{2}$$

Fig. 4 plots individuals' Differenced Earnings (DEs) for each treatment. DEs are most dispersed under *Popularity*. Furthermore, 4, 8 and 29 percent of participants in the *Indifference*, *Random* and *Popularity* treatments, respectively, have *negative* Differenced Earnings. We use these percentages to obtain Result 3.

**Result 3.** When moving away from RSD, the proportion of individuals faring worse is significantly greater under *Popularity* versus *Indifference*; all other comparisons are insignificant. (Two-tailed Fisher's Exact tests at 5% significance.)

4.2. Analysis of behavior

We begin our analysis by checking whether behavior is consistent with the unique symmetric Ex-Post Equilibrium (EPE) induced by *Popularity* and *Indifference* (Theorems 1 and 2). We find that subjects almost never play the EPE (Result 4).

**Result 4.** Only 2 out of 480 and 3 out of 480 reports in the *Popularity* and *Indifference* treatments, respectively, are consistent with the Ex-Post Equilibrium from Theorem 1.

The Truthful strategy (Definition 4.1) is also extremely uncommon in the *Popularity* and *Indifference* treatments (Result 5).

**Definition 4.1.** The Truthful strategy is to state, for all Payoff Functions, the truthful report  $x = \{\{g_1\} > \{g_2\} > \{g_3\} > \{g_4\} > \{g_5\} > \{g_6\}\}$ .

**Table 2**

This table shows the average returns from various strategies against observed opponent play. The first step to obtaining the table entry corresponding to treatment  $T$  and strategy  $S$ ,  $E_T^S$ , is the construction of a 24-by-20 matrix,  $M_T^S$ , where  $M_T^S(i, j)$  is the expected number of Experimental Currency Units that subject  $i$  in treatment  $T$  would receive in Assignment Game  $j$  if she played according to strategy  $S$ , given the observed play of her opponents in game  $j$ . Then,  $E_T^S = [\sum_{i=1}^{24} \sum_{j=1}^{20} M_T^S(i, j)]/480$ ; in other words,  $E_T^S$  is the average entry in  $M_T^S$ . When  $S$  is Truthful,  $i$ 's play is set to the Truthful strategy. When  $S$  is Observed,  $M_T^S(i, j)$  is subject  $i$ 's observed earnings in game  $j$ .

	Truthful	CF1	CF2	CF3	EPE	Observed
<i>Indifference</i>	9.33	38.68	51.30	28.73	10.23	47.41
<i>Popularity</i>	10.05	49.34	60.27	32.56	10.81	46.21

**Result 5.** Only 9 out of 480 and 2 out of 480 reports in the Popularity and Indifference treatments, respectively, are consistent with the Truthful strategy from Definition 4.1.

Popularity and Indifference were designed to induce agents to express cardinal preference information via Cardinal Following (CF) strategies, which are defined below. (See Fig. 3 as well). Empirically, we find that CF reporting is quite common, especially in the *Indifference* treatment (Result 6).

**Definition 4.2.** The Cardinal Following 1 (CF1) strategy is to place good 1 in bin 1. Then, after placing good  $x$  in a bin  $y$ , good  $x + 1$  is placed in bin  $y$  if  $x + 1$  is 2 ECU less in value than  $x$ . Otherwise, good  $x + 1$  is placed in bin  $y + 1$ .

**Definition 4.3.** The Cardinal Following 2 (CF2) strategy is to place good 1 in bin 1. Then, after placing good  $x$  in a bin  $y$ , good  $x + 1$  is placed in bin  $y$  if  $x + 1$  is 2 or 10 ECU less in value than  $x$ . Otherwise, good  $x + 1$  is placed in bin  $y + 1$ .

**Definition 4.4.** The Cardinal Following 3 (CF3) strategy is to place good 1 in bin 1. Then, after placing good  $x$  in a bin  $y$ , good  $x + 1$  is placed in bin  $y$  if  $x + 1$  is 2 or 80 ECU less in value than  $x$ . Otherwise, good  $x + 1$  is placed in bin  $y + 1$ .

**Result 6.** In the *Indifference*, *Popularity* and *Random* treatments, the percentages of CF reports are 61, 34 and 28, respectively.<sup>20</sup> Indifference yields significantly more CF reports than each other mechanism. (Two-Tailed Fisher's Exact tests at 1% significance.)

The prevalence of CF behavior in the *Indifference* and *Popularity* treatments compared to truthful and equilibrium reporting is empirically justified in Table 2: given the observed play of her opponents, the average subject would earn more if she were to play CF1, CF2 or CF3 versus Truthful and EPE.<sup>21</sup> Of these five strategies, CF2 is the most profitable in both treatments. In the *Indifference* treatment, a subject earns 92.42% of the CF2 return. The analogous figure in the *Popularity* treatment is 76.67%.

Although CF reporting is dominated by playing the Truthful strategy under Random, CF behavior is not only observed in the *Random* treatment, CF1 alone is in fact more common than Truthful. CF play under Random may result from altruism. The magnitude of possible other-regarding preferences, however, is limited: less than 14% of reports in the *Random* treatment involve placing a good in the first bin whose value is less than 93 ECU.

While CF reporting is quite common, especially under Indifference, there remains a substantial proportion of choices that it does not explain. To better understand this unexplained behavior, we first ask the fundamental question of whether there are reports that contain "Preference Reversals (PRs)", as defined in Definition 4.5.

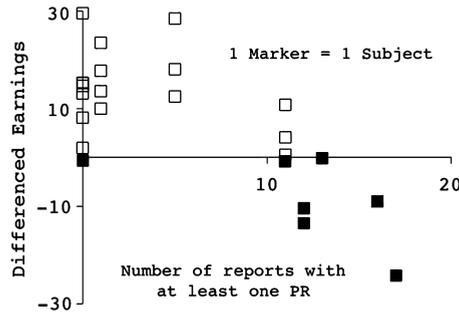
**Definition 4.5.** A report contains a **Preference Reversal** (PR) if it involves goods  $g_i$  and  $g_j$  such that  $g_j$  is strictly lower than  $g_i$  in value but  $g_j$  is reported in an indifference class that is strictly-preferred to the indifference class containing  $g_i$ .

To motivate why an agent may want to make a report containing a PR under Popularity, we present an example in Appendix A.3 where a player in the game induced by Popularity has a unique best-response that contains a PR, giving rise to the following observation.

**Observation 1.** In the game induced by Popularity, if a strategy involves a report that contains a Preference Reversal, the strategy is not necessarily weakly dominated.

<sup>20</sup> Table 4 shows the specific distributions of CF1, CF2 and CF3 reports in each treatment.

<sup>21</sup> We would like to thank an anonymous referee whose suggestions led us to the analysis in Tables 2 and 3 as well as Result 9.



**Fig. 5.** Each marker in the figure is a single subject in the *Popularity* treatment. Subjects with negative Difference Earnings (DEs) are marked as solid black square markers; 6 out of 7 of these individuals make *at least 11* reports containing Preference Reversals (PRs). All subjects with positive DEs make *at most 11* reports that contain PRs.

We are unable to make an analogous observation regarding PRs for Indifference. Under *Random*, reporting one's preferences truthfully is a dominant strategy, leading to few reports containing PRs in the *Random* treatment. We are thus not surprised to find that *Popularity* induces by far the most reports with Preferences Reversals (Result 7).

**Result 7.** Under *Popularity*, *Random* and *Indifference*, 28, 6 and 3 percent of reports contain PRs, respectively. *Popularity* induces significantly more PRs in comparison to *Random* as well as *Indifference*. (Two-Tailed Fisher's Exact tests at 1% significance.)

In light of this result, it seems very unlikely that subjects in the *Popularity* treatment make reports containing PRs accidentally, i.e., as a result of noise. A more plausible explanation is an understanding that such reports have the benefit of possibly improving one's priority in the agent ordering. When a good in a subject's first bin is part of a PR, however, she obtains the lowest-valued good in her first bin more often than not (Result 8).

**Result 8.** Of the reports from the *Popularity* treatment containing PRs, 81% involve placing a good  $g_j$  in bin 1 and a good  $g_i$  in a less-preferred bin despite  $g_i$  being strictly higher in value; 63% of such reports result in the receipt of the lowest-valued good in bin 1.

We (unsurprisingly) find a *negative* relationship between a subject's Differenced Earnings and the number of reports she makes having at least one PR (Fig. 5). In addition, reports with at least one PR yield lower average returns compared to those with none (Result 9).

**Result 9.** In the *Popularity* treatment, the average report containing at least one PR yields a return of 32.39 ECU while the average report with none provides 51.51 ECU.<sup>22</sup>

While these results suggest that subjects make costly mistakes when submitting reports with PRs, it could be that such reports are made in situations where alternative salient strategies would not be any more profitable. In other words, while Table 2 shows that CF2 yields more returns compared to Observed reports in the *Popularity* treatment, deviations from Observed behavior *with one or more PRs* to CF2 may actually cause a *decrease* in returns. The bottom row of Table 3 shows that this is not the case: *any* CF strategy is more profitable compared to the average report with at least one PR. As a result, reports with one or more PRs can indeed be interpreted as costly mistakes. Nevertheless, the returns to reports with at least one PR are still far greater than those from playing the EPE or Truthful strategies. Lastly, note that numbers are quite similar within each column (other than Observed), indicating that the very decision to make a report with at least one PR versus none is not based on the returns from the various strategies listed in the table.

Finally, we consider reports that are not explained by CF strategies and contain no PRs. In the absence of PRs, the Agent Priority Rules under *Indifference* and *Popularity* (defined in Section 2) are the same. We use this common rule to construct a strict weak ordering,  $<$ , where  $x < y$  if an agent stating  $y$  is prioritized, with probability 1, over one stating  $x$ .<sup>23</sup> According to  $<$ , the least element in the set of reports containing no PRs is to state one's preferences truthfully (Definition 4.1).

Using  $<$ , Definition 4.1 and the definitions of the three forms of Cardinal Following behavior, we can organize all reports into Table 4. As far as interpreting the table, the "33" in the *Indifference* row and  $T < y < CF1$  column indicates that, in the

<sup>22</sup> The samples are not independent, so we do not compare them using a Mann-Whitney test. Nonetheless, the difference is meaningful, translating to \$4.78.

<sup>23</sup> Strict weak orderings satisfy irreflexivity, asymmetry, transitivity and transitivity of incomparability.

**Table 3**

This table shows the average returns from various strategies against observed opponent play, organized by reports containing no PRs (top row) or at least one PR (bottom row). The first step to obtaining the entries in the table is to compute two 24-by-20 matrices,  $I_{top}$  and  $I_{bottom}$ . Entry  $I_{top}(i, j)$  equals 1 if subject  $i$ 's Observed report in Assignment Game  $j$  contains no PRs; otherwise,  $I_{top}(i, j) = 0$ . Similarly,  $I_{bottom}(i, j) = 1$  if subject  $i$ 's Observed report in Assignment Game  $j$  contains at least one PR and  $I_{bottom}(i, j) = 0$  otherwise. Note that  $[\sum_{i=1}^{24} \sum_{j=1}^{20} I_{top}(i, j)] = 347$  and  $[\sum_{i=1}^{24} \sum_{j=1}^{20} I_{bottom}(i, j)] = 133$ . Then, to obtain the top and bottom table entries corresponding to strategy  $S$ ,  $E_{top}^S$  and  $E_{bottom}^S$ , we construct a 24-by-20 matrix,  $M^S$ , where  $M^S(i, j)$  is the expected number of Experimental Currency Units that subject  $i$  would receive in Assignment Game  $j$  if she played according to strategy  $S$ , given the observed play of her opponents in game  $j$ . Then,  $E_{top}^S = [\sum_{i=1}^{24} \sum_{j=1}^{20} M^S(i, j)I_{top}(i, j)]/347$  and  $E_{bottom}^S = [\sum_{i=1}^{24} \sum_{j=1}^{20} M^S(i, j)I_{bottom}(i, j)]/133$ . When  $S$  is Truthful,  $i$ 's play is set to the Truthful strategy. When  $S$  is Observed,  $M^S(i, j)$  is subject  $i$ 's observed earnings in game  $j$ .

	Truthful	CF1	CF2	CF3	EPE	Observed
Indifference	9.33	38.68	51.30	28.73	10.23	47.41
Popularity	10.05	49.34	60.27	32.56	10.81	46.21

**Table 4**

Each report ( $r$ ) is categorized across all three treatments. Bolded figures indicate the treatment that generates the most reports for a given report category.

	Reports with PRs	$r = \text{Truthful (T)}$	$T < r < \text{CF1}$	$r = \text{CF1}$	$\text{CF1} < r < \text{CF2}$	$r = \text{CF2}$	$\text{CF2} < r < \text{CF3}$	$r = \text{CF3}$	$\text{CF3} < r$	total	$r = \text{EPE}$
Popularity	<b>133</b>	9	70	120	27	24	75	19	3	480	2
Indifference	12	2	33	<b>139</b>	<b>31</b>	<b>109</b>	<b>95</b>	<b>45</b>	<b>14</b>	480	<b>3</b>
Random	29	<b>102</b>	<b>186</b>	126	6	10	21	0	0	480	0

Reports with no Preference Reversals (PRs)

Indifference treatment, there are 33 reports such that a deviation to the Truthful (CF1) strategy would yield the reporter a weakly later (earlier) position in the agent ordering.<sup>24</sup> In each column, the largest number is bolded.

Table 4 reiterates some of our previous findings regarding reports containing PRs, CF reporting and EPE play. In all treatments, we see that CF1 is more common than CF2 and CF2 is more common than CF3. Furthermore, when we purely focus on reports containing no PRs, it is easy to see that Indifference generates more “binning” than Random, according to  $\text{CF1} <$ . In each of the two left-most columns of reports with no PRs, the Random treatment accounts for most of the behavior.<sup>25</sup> For each of the six remaining categories of reports containing no PRs, Indifference explains the most behavior.

4.3. Summary of results

In terms of aggregate welfare, all three of the algorithms we test (Popularity, Indifference and Random) show improvements over RSD. Thus, simply allowing indifference can yield welfare improvements, even if true preferences are strict. Nonetheless, the aggregate earnings are significantly higher under Indifference and Popularity compared to Random, showing that even more welfare gains can be obtained by actually incentivizing agents to express indifference in a very simple and user-friendly manner. On an individual level, however, Indifference outperforms Popularity.

In terms of behavior, over half of the reports under Indifference are consistent with the mechanism’s intended Cardinal Following strategies. Thus, these reports seem to be focal when subjects are faced with this mechanism. While CF behavior is also present under Popularity, there are also many reports that contain Preference Reversals (PRs). Thus, Popularity does not focalize the same type of reports as does Indifference. We find that reports containing PRs tend to be unprofitable and that subjects make such reports with different propensities, driving the inequality in earnings we observe under Popularity.

5. Related literature and discussion

There is a previous literature that experimentally compares the welfare properties of strategy-proof and manipulable assignment mechanisms. Klijn et al. (2013) and Lien et al. (2015) report the manipulable Boston mechanism having

<sup>24</sup> The change in agent ordering is “weak” because the deviation could lead to no change. This would be true, for instance, if all opposition played the EPE and  $y$  was such that  $T < y < \text{CF1}$ .

<sup>25</sup> The Dominant-Strategy Equilibrium (DSE) in Random is the truthful strategy, thus explaining the left-most column of reports with no PRs.

mixed success against strategy-proof alternatives.<sup>26</sup> Chen and Sönmez (2006) find that the manipulable Boston mechanism performs worse than two strategy-proof alternatives in a school choice environment. Calsamiglia et al. (2010) study the impacts of constraining the length of preference lists agents may submit. They show that such constraints interfere with strategy-proofness and ultimately lower welfare. While Abdulkadiroğlu et al. (2011) show that the Boston mechanism may outperform strategy-proof mechanisms like DA in equilibrium,<sup>27</sup> Featherstone and Niederle (2016) find that experimental subjects have difficulties finding non-truthtelling equilibria under the Boston mechanism. Other manipulable mechanisms may not perform well either: Hugh-Jones et al. (2014) find that while the probabilistic serial mechanism of Bogomolnaia and Moulin (2001) has desirable welfare properties in theory, lab subjects often fail to misreport their preferences when it can be beneficial yet misstate them when there is no benefit of doing so; in other words, the welfare-improving strategies are not focal.

Our paper differs from most of this literature by showing that welfare gains from non-truthful reporting are indeed possible in practice. We think our success comes from the fact that we have designed a new mechanism that explicitly focalizes (non-truthful) welfare-improving strategies, as opposed to exclusively studying manipulable mechanisms that are common in the field. While understanding the properties of existing field mechanisms is inarguably important, their popularity may in some sense be a “historical accident,” and the behavior they induce may not have been carefully designed or tested before implementation.

Another possibility for welfare improvements is to attempt to elicit cardinal preferences directly. Such an approach is taken in Hylland and Zeckhauser (1979), who provide an algorithm that is ex-ante Pareto optimal.<sup>28</sup> Despite the mechanism being proposed almost forty years ago, it is not (to our knowledge) used anywhere in practice. This conforms with the common intuition in practical market design that ordinal preferences are easier for real-world agents to report than cardinal ones (for example, in an experiment assigning MBA students to courses at Wharton, Budish and Kessler (2014) state that subjects “had particular difficulty with reporting cardinal preference intensity information”).<sup>29</sup> With respect to cardinal-languages in the field, “fake-money” auctions are what we have seen the most. The problem with such auctions is that when they end, agents may be left with large budgets of useless currency (Sönmez and Ünver, 2010a; Budish and Kessler, 2014). Solutions have been suggested, such as the approximate competitive equilibrium from equal incomes (ACEEI) mechanism used by Budish (2011) to make multi-unit assignments. When applied to our single-unit environment, however, ACEEI actually collapses to RSD.<sup>30</sup> Extending the ideas of this paper to multi-unit assignment is an interesting direction for future work.

The Popularity and Indifference mechanisms discussed in this paper are similar to the Choice Augmenting Deferred Acceptance (CADA) mechanism from Abdulkadiroğlu et al. (2015) in that all three mechanisms allow agents to affect their priorities via their reports. We describe CADA using a school choice setting, as in Abdulkadiroğlu et al. (2015). In CADA, a student states her ordinal rankings over schools; she has a dominant strategy of doing so truthfully. In addition, she may choose a school to “augment” which serves as a tie-breaker in the following sense: if a school  $s$  places students  $i$  and  $j$  in the same priority class, it is randomly determined who ultimately receives the higher priority at  $s$  when CADA ultimately runs the Deferred Acceptance (DA) mechanism (Gale and Shapley, 1962) unless precisely one of these students augments  $s$ , in which case this student obtains a strictly higher priority compared to the other at  $s$ . Augmenting a school is thus a way that a student can express some information on her cardinal preferences.

CADA is an interesting mechanism that is likely to work very well in some settings. For example, if each agent  $i$  similarly values schools 1 through  $s_i$ , strongly prefers  $s_i$  to  $s_i + 1$  and similarly values schools  $s_i + 1$  through  $n$  for some  $s_i \in \{1, \dots, n - 1\}$ , then agents will be able to essentially communicate their cardinal preferences under CADA by augmenting  $s_i$ . When agents' preferences are not so dichotomous, efficient allocations from a social planner's perspective may be easier to attain via Popularity or Indifference. For instance, in our experimental environment, suppose the cardinal values for goods 1 through 6 held by agents  $a$  and  $b$  are given by  $v_a = \{99, 97, 17, 15, 5, 3\}$  and  $v_b = \{99, 97, 17, 7, 5, 3\}$ , respectively. Under CADA, it seems reasonable for  $a$  and  $b$  to report both the same ordinal preferences and also express the same cardinal message by augmenting good 2. It is possible, however, that neither  $a$  nor  $b$  receives good 1 or 2. In this case, further

<sup>26</sup> Klijn et al. (2013) find that the strategy-proof deferred acceptance (Gale and Shapley, 1962) outperforms the manipulable Boston mechanism in terms of efficiency unless agents are only permitted to submit constrained preference lists. Lien et al. (2015) show that the Boston mechanism can actually be more efficient than a serial dictatorship in practice in a school admissions problem where students have to submit preferences before colleges learn the students' exam scores that serve as noisy signals of the students' academic abilities.

<sup>27</sup> See also Miralles (2008) and Troyan (2012).

<sup>28</sup> Ex-ante Pareto optimality means that, given submitted preferences, the probabilities that agents receive over various goods are efficient, i.e., for any trade in shares that strictly increased some individual's expected utility, another individual's expected utility must be strictly reduced. Zhou (1990) proves an impossibility result that no ex-ante Pareto optimal and symmetric mechanism is strategy-proof.

<sup>29</sup> Motivated by such observations, Carroll (2011) studies the question of when it is without loss of generality for a designer to restrict to ordinal mechanisms.

<sup>30</sup> The ACEEI mechanism of Budish (2011) is targeted for solving multi-unit assignment problems in a way that ensures fairness and approximate efficiency ex-post. While in practice agents may be asked for “cardinal” values of all objects (see Budish and Kessler, 2014), this information is actually converted into an ordinal preference relation for purposes of running the algorithm. It would be prohibitively complex and time-consuming to ask agents to report an ordinal preference relation over all possible bundles (e.g., over all possible course schedules in a business school) when there is multi-unit demand; asking for cardinal values is only used to simplify the reporting language (with the restriction that some preferences may not be expressible). ACEEI is still an ordinal mechanism, and in the case of single-unit assignment, ACEEI collapses to RSD. Thus, while ACEEI seems to be a good solution for multi-unit assignment, it is designed to address issues different from those that we study in this paper.

information differentiating  $a$  and  $b$  would be useful since CADA may allocate goods 3 and 4 to  $a$  and  $b$ , respectively, but a social planner would prefer  $a$  and  $b$  to receive 4 and 3, respectively. Under Indifference or Popularity,  $a$  and  $b$  could report  $\{g_1, g_2\} > \{g_3, g_4\} > \{g_5, g_6\}$  and  $\{g_1, g_2\} > \{g_3\} > \{g_4, g_5, g_6\}$ , respectively, and likely induce the planner's assignment. Abdulkadiroğlu et al. (2015) nonetheless identify multiple potential benefits of CADA, and testing its potential impact is an interesting question for future work, as it has the potential to perform extremely well in some settings.

To summarize, this paper is motivated by a very common problem faced by many institutions, that of assigning indivisible resources to individuals without using monetary transfers. Canonical strategy-proof mechanisms such as RSD may have limited impact, especially in settings where ordinal preferences for the goods are correlated, because they do not provide channels through which individuals can express cardinal preference information. We investigate new mechanisms that provide agents with such channels. While equilibrium predictions suggest that they should not outperform RSD, we find that they do so empirically by focalizing particular strategies which are welfare-improving.

We see this paper as contributing to a rapidly growing “behavioral market design” literature, which we think of as market design for agents with strategic limitations. In a sense, this literature is quite old; for years, matching theorists have required that their designed mechanisms be strategy-proof, i.e., give individuals dominant strategies of reporting truthfully. However, strong incentive constraints may hinder efficiency by not allowing agents to transmit any information on their cardinal preferences. In some settings, it may be possible to relax these constraints while still ensuring that agents behave in predictable (and desirable) ways. In this paper, we do not aim to focalize truth-telling, but rather, non-truthful cardinal-following reports where individuals state indifference across goods that are similar in value. In the mechanism we construct, 61% of play is cardinal-following, and cardinal welfare is significantly higher than under the standard strategy-proof alternative. As economic theory and policy evolve from a world of perfectly rational agents to one where people have cognitive limitations, it is important that future research in market design continues to head in this direction as well.

## Appendix A

### A.1. Proof of Theorem 1

All arguments are equivalent for both the Popularity and Indifference mechanisms, and so we will not distinguish between the two. We use the Lemma 1 to prove Theorem 1.

**Lemma 1.** *Let  $x$  be a message profile such that  $x_i \neq x_i^{AbL}$  and  $x_{-i} = x_{-i}^{AbL}$ . Then, when  $x$  is submitted to Indifference or Popularity,  $i$  receives good  $n$  with probability 1.*

**Proof.** Let  $x$  be a submitted message profile where  $x_i \neq x_i^{AbL}$  and  $x_{-i} = x_{-i}^{AbL}$ . If  $x_i$  has good  $n$  in its first bin, every allocation that gives every agent a good in their first bin must give  $n$  to  $i$ . If  $x_i$  does not have  $n$  in its first bin, then its first bin must contain a strict subset of the best  $n - 1$  goods. (Otherwise,  $x_i = x_i^{AbL}$ .) In this case,  $i$  is ordered last and receives  $n$  with certainty.  $\square$

To prove Theorem 1, the following must hold for all  $i, v_i \in V_i$ :

$$u_i(v_i, x_i^{AbL}, x_{-i}^{AbL}) \geq u_i(v_i, x'_i, x_{-i}^{AbL}).$$

This immediately follows by Lemma 1, however, since  $u_i(v_i, x'_i, x_{-i}^{AbL}) = v_i(n) = 0 < (1/n) \times \sum_{k=1}^n v_i(k) = u_i(v_i, x_i^{AbL}, x_{-i}^{AbL})$  for any  $x'_i \neq x_i^{AbL}$  and regardless of  $i$ 's type,  $v_i$ .

### A.2. Proof of Theorem 2

All arguments are equivalent for both the Popularity and Indifference mechanisms, and so we will not distinguish between the two. We use the Lemmas 1 and 2 to prove Theorem 2.

**Lemma 2.** *Let  $x$  be a message profile such that  $x_i = x_i^{AbL}$  and  $x_{-i} \neq x_{-i}^{AbL}$ . Then, when  $x$  is submitted to Indifference or Popularity,  $i$  receives good  $n$  with probability 0.*

**Proof.** Since  $x_{-i} \neq x_{-i}^{AbL}$ , there exists some  $j \neq i$  such that  $x_j \neq x_j^{AbL}$ . If  $i$  receives good  $n$  then  $j$  receives some good  $t < n$ ; it is Pareto improving with respect to their designated bins if  $i$  and  $j$  exchange their goods with each other. Thus, Indifference and Popularity cannot assign good  $n$  to  $i$ . If  $x_j$  does not have  $n$  in its first bin, then its first bin must contain a strict subset of goods 1 through  $n - 1$ . (Otherwise,  $x_j = x_j^{AbL}$ .) Thus,  $i$  is ordered before  $j$ . Since  $i$  is *not* ordered last,  $i$  is not assigned  $n$ .  $\square$

To prove Theorem 2, let  $s = (s_1, \dots, s_n)$  be a symmetric ex-post equilibrium. Consider some vector of cardinal utilities  $v^*$  such that  $v^*(n - 1) > \bar{v}$ , where  $\bar{v} = (1/n) \times \sum_{k=1}^n v^*(k)$ . Consider the type profile where  $v_j = v^*$  for all agents  $j \in N$ . By symmetry, each agent plays the same strategy,  $s_i(v^*) = s_j(v^*)$  for all  $i, j$ , and so each receives each good with a  $1/n$  probability; hence,  $u_i(v^*, s_i(v^*), s_{-i}(v_{-i}^*)) = \bar{v} < v^*(n - 1)$ .

Assume that  $s_i(v^*) \neq x_i^{AbL}$ . Note that if agent  $i$  deviates and reports  $x_i^{AbL}$ , by Lemma 2, her probability of receiving good  $n$  is 0, and so

$$u_i(v^*, x_i^{AbL}, s_{-i}(v_{-i}^*)) \geq v^*(n - 1) > \bar{v}_i = u_i(v^*, s_i(v^*), s_{-i}(v_{-i}^*)).$$

In other words, this deviation guarantees the receipt of good  $n - 1$  or better and thus is strictly profitable. Thus, if  $v^*(n - 1) > \bar{v}_i$ , then  $s_i(v^*) = x_i^{AbL}$ ; by symmetry,  $s_j(v^*) = x_j^{AbL}$  for all  $j \in N$ .

Now suppose that there exists some  $v_i \in V_i$  such that  $s_i(v_i) \neq x_i^{AbL}$ . Consider the type profile such that  $v_j = v^*$  for all  $j \neq i$ . By the previous paragraph, we have  $s_j(v_j) = x_j^{AbL}$  for all  $j \neq i$ . Then,  $u_i(v_i, s_i(v_i), x_{-i}^{AbL}) = v_i(n) = 0$  (by Lemma 1), while  $u_i(v_i, x_i^{AbL}, x_{-i}^{AbL}) = \bar{v}_i > 0$ , and so  $x_i^{AbL}$  is a profitable deviation. Hence,  $s_i(v_i) = x_i^{AbL}$  for all  $v_i \in V_i$ , i.e.,  $s_i = s_i^{AbL}$ . The same argument applies for all agents, and so  $s = s^{AbL}$ . This completes the proof.

### A.3. Example motivating preference reversals under popularity

Let  $p = (p_1, \dots, p_6)$  denote the portfolio of shares over goods  $g_1, \dots, g_6$  that a player  $i$  receives from a set of submitted reports. Let  $v(p)$  denote  $i$ 's expected value of portfolio  $p$  (in ECU), given  $i$ 's valuations for the goods. Suppose that  $v((1/3, 0, 2/3, 0, 0, 0)) > v((1/4, 1/4, 1/2, 0, 0, 0))$ . Suppose  $i$  states  $r = \{g_1, g_3\} > \{g_2, g_4, g_5, g_6\}$ , an agent  $j$  states  $r_j = \{g_1, g_2, g_3\} > \{g_4, g_5, g_6\}$  and the remaining four state  $r_4 = \{g_1, g_2\} > \{g_3, g_4, g_5, g_6\}$ . Under this profile,  $p = (1/3, 0, 2/3, 0, 0, 0)$ . If  $i$  deviates to a report such that her first bin is  $\{g_1, g_2, g_3\}$ , then  $p$  becomes  $(1/4, 1/4, 1/2, 0, 0, 0)$ . If  $i$  deviates to  $r_4$ , then  $j$  receives  $g_3$  with probability 1 and, by symmetry, all other agents (including  $i$ ) have equal chances of acquiring the remaining goods. For all possible valuations that  $i$  could have for the goods,  $v((1/4, 1/4, 1/2, 0, 0, 0)) > v((1/5, 1/5, 0, 1/5, 1/5, 1/5))$ . Lastly, if  $i$  deviates to a report not equal to  $r_4$  such that her first bin is  $\{g_1, g_2\}$ , is  $\{g_1\}$  or contains  $g_4, g_5$  or  $g_6$ , then the probability that  $i$  is assigned  $g_1, g_2$  or  $g_3$  is 0. Thus,  $i$  is strictly better off stating  $r$  (which contains a PR) versus any other report that contains no PRs.

If Indifference were used instead of Popularity, if  $i$  states  $r$  or  $r_4$ , agent  $j$  is ordered first and all other agents are ordered at random, thus giving  $i$  a  $3/5$  chance of being one of the three agents having the lowest priorities. Conditional on this occurrence, she has equal chances of obtaining  $g_4, g_5$  and  $g_6$ . Conditional on being the agent prioritized second or third,  $i$ 's respective chances of obtaining goods  $g_1, g_2$  and  $g_3$  are  $1/3, 0$  and  $2/3$  under  $r$  while they are  $1/2, 1/2$  and  $0$  under  $r_4$ . Thus, reporting  $r_4$  is strictly better for  $i$  than reporting  $r$  (under Indifference), irrespective of her cardinal preferences.

### A.4. Experimental instructions

Instructions were presented using a PowerPoint presentation. We first show the slides shown that were common to all treatments, then we show those that are treatment-specific, i.e., that define the particular mechanism for the given treatment. After going through the slide show as a group, subjects received printouts of the instructions to use for reference as they made their decisions. Subjects also received pens and scratch paper but no calculators.

## Assigning Students To Trips Abroad

Version 2

Welcome! My name is Dan Fragiadakis

### Instructions common to all treatments

We will soon start this presentation of the instructions for today's study. You may take notes if you wish, but you will receive printouts of this presentation after we go through it together, which you can keep for reference during the study.

#### Introduction (true story)

- The Stanford Graduate School of Business (GSB) sends its students to foreign countries
- Trips are 2 weeks long
- Students meet policymakers, CEOs...etc
- Destinations include China, Brazil, South Africa... and ~20 others
- Purpose of this study: How can we help the GSB give students their favorite trips?

#### As real as it gets

- Even though we can't fly you to Hong Kong,
  - 1) you will walk out today with CASH
  - 2) your decisions will be used to help improve the GSB's process of assigning students to trips
- This study has a direct impact here at Stanford, so we ask that you take this experiment seriously

## What do we mean by serious?

- We ask that you follow a few ground rules, PLEASE
- 1) No Phones
  - 2) No Talking or Exclaiming Out Loud
  - 3) No Looking at Each Others Screens
  - 4) If you have any questions, please do not shout them out; instead, please raise your hand

## Preferences over trips (continued)

We pay you according to the trip you're assigned

The amounts you are paid for the trips are PRIVATE; no other subjects see what you are paid

In the real world, GSB students' preferences may be correlated; in other words, certain trips are likely to be "popular" and others "unpopular"

## Payments for trips

- Here are some values for a student, Bob

Trip	1	2	3	4	5	6
Value	99	97	17	15	5	3

- Here are some values for a student, Kim

Trip	1	2	3	4	5	6
Value	99	19	17	7	5	3

- Bob and Kim both like Trip 1 > Trip 2 > ... > Trip 6
- Everyone always values Trip 1 at 99
- Everyone always values Trip 6 at 3

## What YOU DO (continued)

- You must put each trip into a bin; in other words, you must place each trip somewhere
- You can place as many or as few trips into a bin as you like
- You can't "skip" bins. In other words, if you try to click "continue" and you've placed trips in Bin A and Bin C but NOT in Bin B, you've "skipped" Bin B.

This is like trying to tell the algorithm that you have some 1<sup>st</sup> choices (in Bin A) and some 3<sup>rd</sup> choices (in Bin C) but no 2<sup>nd</sup> choices (since you've left Bin B empty).

We think this doesn't make much sense, so if you place a trip in Bin A and a trip in Bin C, you *must* have a trip in Bin B before you continue.

- Even though you can't "skip" bins, you don't have to use all the bins; you can use between 1 and 6 bins

## Preferences over trips

- GSB students have preferences over going on the different trips:

- 1) Bob likes Italy > France > Japan > ...
- 2) Kim likes France > Japan > Mexico > ...

- In this experiment, there are 6 trips (Trip 1, Trip 2 ... Trip 6)

## Preferences over trips (continued)

To simplify matters in this experiment, we construct your preferences such that EVERYBODY

- Has Trip 1 as their favorite trip (i.e. gets paid the most for Trip 1)
- Has Trip 2 as their 2<sup>nd</sup> favorite trip
- Has Trip 3 as their 3<sup>rd</sup> favorite trip
- ...etc

Thus, in terms of ranking the trips in a list from most to least favorite, you all have the same ranking.

HOWEVER, you and your group members' preferences will differ in terms of *how much* you like the trips; the next slide has an example

## What YOU DO

- You will act as a GSB student trying to get a trip
- You will "tell us which trip you want" by placing trips into "Bins"
- There are 6 Bins: Bin A, Bin B ... Bin F
- The algorithm first tries to get you a trip from Bin A. It interprets what you place into Bin A as your favorite choice(s).
- It views you as indifferent between everything *within* Bin A
- If it can't get you a trip from Bin A, it tries to get you a trip from Bin B. It interprets what you place into Bin B as your 2<sup>nd</sup> favorite choice(s).
- It views you as indifferent between everything *within* Bin B
- If it can't get you a trip from Bin B, it tries to get you a trip from Bin C... etc.

## Groups

- You will be divided into groups of 6 people
- (There may be just 1 group in your session)
- Each group is independent from all other groups
- Within each group, each trip has 1 spot. Therefore, the 6 members in a group each get a unique trip.

## What the algorithm does/doesn't know

- The algorithm DOES NOT KNOW what you are paid for the trips (*remember, this is your private information; in reality, only you know what you like and how much you like it*)
- The algorithm ONLY KNOWS what trips you place into what bins (*because this is all you can tell the algorithm*)
- *Remember:* It interprets your Bin A trip(s) as your favorite(s), then your Bin B trip(s) as your second favorite(s)... etc. and it views you as indifferent towards all trips *within* the same bin

## (Slides Explaining how mechanism from treatment works)

## 20 INDEPENDENT periods

In each...

You will play this *trip assignment game*.

Everyone gets new values for the trips

Everyone puts the trips into bins

Everyone gets a trip

## Payments

At the end of the experiment, we randomly select one of the 20 periods and look at the value of the trip you were assigned in that period; suppose that value is  $V$

We take  $V$  and divide it by 4 to get an amount in dollars

Thus, if  $V = 97$ , this becomes  $97/4 = \$24.25$ . You will be paid this.

We also pay you \$5 for showing up, and \$5 for completing the study

Thus, in this example, you would go home with \$34.25.

## Time Limits

- In each period, you'll have 60 seconds to place trips in bins (please think about your decisions carefully; there is no benefit to racing through your choices since we must wait for everyone else to make their choices before the experiment can proceed)
- At the end of 60 seconds, we do not automatically move you to the next screen, but we do urge you to reach a decision so as to not hold up the experiment; we promised an experiment lasting 75 minutes to you and everyone else 😊
- You may place trips into/out of bins as you like, but once you click "continue", you CANNOT go back and change your choices

## Feedback

Before you start a period, we will provide you with information about values, choices, and outcomes from previous periods; in other words, you'll see your *history* in a set of History Tables

Specifically, you'll see your

1. Past trip values
2. Your past choices of which trips you placed into which bins
3. The trips you were assigned

Note: for 2. above, to show you that you placed Trips 1, 4, and 6 in Bin B, for instance, we write 146 under the "Bin B" column of your History Table. If there is a "0" in the "Bin C" column, it means that you placed no trips in Bin C.

## Time for some action...

- We're now ready to begin the experiment!
- We will pass out a copy of the presentation for you to keep for reference
- If you have any questions now or at any time, please raise your hand 😊

## (Slides summarizing treatment)

### Additional slides for the Popularity treatment

#### Creating the Student Ranking

- The algorithm calculates “Demand” for each trip
- Demand = # of people who put the trip in Bin A

	Trip	Demand
Least Demanded →	Trip 1	0
	Trip 2	1
	Trip 6	1
	Trip 4	4
	Trip 3	5
Most Demanded →	Trip 5	6

ties broken RANDOMLY

#### Creating the Student Ranking: example 2

- We use “Demand” to get a “Student Ranking”
- How do we order Bob and Kim?
- find Bob’s *least demanded trip from Bin A*

Trip	Demand
Trip 1	0
Trip 2	1
Trip 6	2
Trip 4	4
Trip 3	5
Trip 5	5

Bob’s Bin A = {Trip 6, Trip 5, Trip 4}

Kim’s Bin A = {Trip 6, Trip 3}

Bob has least demanded trip, so goes higher on the Student Ranking

#### Creating the Student Ranking: example 4

- We use “Demand” to get a “Student Ranking”
- How do we order Bob and Kim?
- find Bob’s *least demanded trip from Bin A*

Trip	Demand
Trip 1	0
Trip 2	2
Trip 6	3
Trip 4	3
Trip 3	4
Trip 5	5

Bob’s Bin A = {Trip 2, Trip 4}

Kim’s Bin A = {Trip 2, Trip 4}

Both students “run out” of trips to compare in Bin A, so we move to Bin B and try to compare them using Bin B contents

### Creating an Assignment: Overview

- After you make your choices (i.e. place the trips in bins), the algorithm uses these decisions to create a Student Ranking (i.e. an ordered list of you and your group members: someone will be 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>...etc in the Student Ranking)
- According to the Student Ranking, the algorithm gives people their most preferred trips
- Let’s see how we use students’ decisions to determine which one of them gets ranked higher on the Student Ranking

#### Creating the Student Ranking: example 1

- We use “Demand” to get a “Student Ranking”
- How do we order Bob and Kim?
- find Bob’s *least demanded trip from Bin A*

Trip	Demand
Trip 1	0
Trip 2	1
Trip 6	1
Trip 4	4
Trip 3	5
Trip 5	6

Bob’s Bin A = {Trip 5, Trip 6}

Kim’s Bin A = {Trip 5, Trip 3, Trip 4}

Bob has least demanded trip, so goes higher on the Student Ranking

#### Creating the Student Ranking: example 3

- We use “Demand” to get a “Student Ranking”
- How do we order Bob and Kim?
- find Bob’s *least demanded trip from Bin A*

Trip	Demand
Trip 1	0
Trip 2	1
Trip 6	2
Trip 4	4
Trip 3	4
Trip 5	5

Bob’s Bin A = {Trip 6, Trip 5, Trip 4}

Kim’s Bin A = {Trip 6}

Kim “runs out” of trips to compare, so Bob automatically goes higher on the Student Ranking

### Creating the Student Ranking

- If we must compare two students using their Bin B contents, the way we compare them is identical to the way we compare students’ Bin A
- In other words, we look at their lowest demanded trips and see which student has the lesser demanded trip...
- ...crossing out trips that are ties
- If we end up crossing out everything in Bin B for both students, we compare them based on Bin C... etc
- Note: Even though we may compare students’ Bin B contents, we still use the original Demand list that was computed using Bin A contents

## Creating the Student Ranking (summary)

- We start by trying to rank students according to Bin A
- If we can't rank two students based on Bin A, we try to rank them based on Bin B. If we can't...etc.
- As we just saw, you get higher on the student ranking by putting trips with low demand in your more preferred bins
- Although it took awhile to go through the Student Ranking, the next slide illustrates how important the ranking is

## The Student Rank (SR)

The algorithm first computes the SR. In this example, it turns out we only need to look at Bin A and Bin B to find the SR.

Trip	Demand	SR	Bin A	Bin B
Trip 1	0	1 <sup>st</sup> : Jane	Trip 3, Trip 5	Trip 4
Trip 4	0	2 <sup>nd</sup> : Jeff	Trip 2	Trip 1
Trip 6	0	3 <sup>rd</sup> : Bob	Trip 2	Trip 3
Trip 5	1	4 <sup>th</sup> : Sarah	Trip 3	Trip 4
Trip 2	2	5 <sup>th</sup> : Alex	Trip 3	Trip 6
Trip 3	4	6 <sup>th</sup> : Rachel	Trip 3	Trip 2

## Making assignments: round 1

The algorithm first gives as many students as possible something from Bin A, possibly "skipping" students if your Bin A trip is taken

Trip	Demand	SR	Bin A
Trip 3	4	1 <sup>st</sup> : Jane	Trip 3, Trip 5
Trip 2	2	2 <sup>nd</sup> : Jeff	Trip 2
Trip 5	1	3 <sup>rd</sup> : Bob	Trip 2
Trip 6	0	4 <sup>th</sup> : Sarah	Trip 3
Trip 4	0	5 <sup>th</sup> : Alex	Trip 3
Trip 1	0	6 <sup>th</sup> : Rachel	Trip 3

**NOTE:** when the algorithm couldn't give Trip 2 to Bob, it did NOT consider Bob's Bin B. Instead, it moved to Sarah.

## Making assignments: round 2

Recall that Jane, Jeff and Sarah have Trips 5, 2 and 3

Now, we consider the remaining students one-at-a-time: if we can't give you something from Bin B, we try to give you something from Bin C, if we can't, we move to Bin D...etc

In other words, in Round 1, we skipped students when we couldn't give them something from Bin A, but in Round 2, we do NOT skip, but rather, consider your NEXT preferred bin(s)

SR	Bin B	Bin C	Bin D
3 <sup>rd</sup> : Bob	Trip 3	Trip 6	Trip 5
5 <sup>th</sup> : Alex	Trip 6	Trip 5, Trip 2	Trip 1, Trip 4
6 <sup>th</sup> : Rachel	Trip 2	Trip 1	Trip 4

## Making assignments: summary

We order students according to the Student Ranking

Round 1: We give you something from Bin A if possible. If not, we skip you until round 2, and try to give the next person on the Student Ranking something from Bin A.

Round 2: If you didn't get a trip in Round 1, we give you a trip in Bin B if possible. If we can't, we try to give you a trip in Bin C. If we can't, we go for Bin D... etc. We do not skip you. We only go to the next person in the Student Ranking once you get a trip.

## The final assignment

SR	Bin A	Bin B	Bin C	Bin D
1 <sup>st</sup> : Jane	Trip 3, Trip 5	Trip 6	Trip 1	Trip 4
2 <sup>nd</sup> : Jeff	Trip 2	Trip 5, Trip 3	Trip 4, Trip 1	Trip 6
3 <sup>rd</sup> : Bob	Trip 2	Trip 3	Trip 6	Trip 5
4 <sup>th</sup> : Sarah	Trip 3	Trip 1	Trip 4	Trip 2, Trip 6
5 <sup>th</sup> : Alex	Trip 3	Trip 6	Trip 5, Trip 2	Trip 1, Trip 4
6 <sup>th</sup> : Rachel	Trip 3	Trip 2	Trip 1	Trip 4

## The Benefits and Costs of Bundling

The more trips you have in earlier bins, *especially if those trips are under-demanded*, the greater your chance of appearing earlier in the Student Ranking

**"Bundling" multiple trips together has two BENEFITS**

1. Going earlier and possibly getting a trip from a more preferred bin
2. Allowing the algorithm to swap you around *within* your bin after you get a trip, which accommodates others

**However, "bundling" has the COST not necessarily getting your favorite trip within a bin**

1. Either because you the algorithm chose a random trip from your bin, or
2. Because the algorithm swapped you around *within* your bin after you get a trip to accommodate others

## Summary

- This experiment is 20 periods
- You will see your private values for the trips
- You submit choices (place the trips into bins)
  - Using subjects' Bin A, we find "Demand" of trips
  - We create the Student Ranking by placing students earlier who prefer LOWER DEMANDED trips
  - We go through the Student Ranking in two rounds:
    - Round 1: give as many people a trip from Bin A as possible, possibly skipping you if we can't give you a trip from Bin A.
    - Round 2: give remaining people their favorite trips. Start with person highest in rank and try to give her something from Bin B. If it can't be done, try to give her something from Bin C. If not, then Bin D... etc. In other words, we DO NOT SKIP to the next person on the Student Ranking list until she is given a trip.

### Additional slides for the Indifference treatment

## Creating the Student Ranking

- We compare students according to the number of trips in Bin A
- Whoever has more trips in Bin A goes earlier on the Student Ranking
- How do we order Bob and Kim?
- Bob’s Bin A = {Trip 5, Trip 6}
- Kim’s Bin A = {Trip 5, Trip 3, Trip 4}
- **Kim has more trips in Bin A, so she goes higher than Bob on the Student Ranking**

## Creating the Student Ranking

- If students have the same number of trips in Bin A and the same number of trips in Bin B, we compare them according to Bin C.
- Whoever has more trips in Bin C goes earlier.
- If they have the same number of trips in Bin C, we move to Bin D... etc.

### Example Student Rank

SR	Bin A	Bin B	Bin C
1 <sup>st</sup> : Jane	Trip 3, Trip 5	Trip 6	Trip 1
2 <sup>nd</sup> : Jeff	Trip 2	Trip 5, Trip 3	Trip 4, Trip 1
3 <sup>rd</sup> : Bob	Trip 2	Trip 3, Trip 4	Trip 6
4 <sup>th</sup> : Sarah	Trip 3	Trip 1	Trip 4, Trip 2, Trip 6
5 <sup>th</sup> : Alex	Trip 3	Trip 5	Trip 6, Trip 1
6 <sup>th</sup> : Rachel	Trip 3	Trip 2	Trip 6, Trip 1

- Jane beats everyone based on Bin A
- Jeff and Bob are next because they have more trips in Bin B than Sarah, Alex and Rachel; Jeff goes before Bob because of Bin C
- Sarah, Alex and Rachel are comparable using Bin C

## Creating an Assignment: Overview

- After you make your choices (i.e. place the trips in bins), the algorithm uses these decisions to create a Student Ranking (i.e. an ordered list of you and your group members: someone will be 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>...etc in the Student Ranking)
- According to the Student Ranking, the algorithm gives people their most preferred trips
- Let’s see how we use students’ decisions to determine which one of them gets ranked higher on the Student Ranking

## Creating the Student Ranking

- If students have the same number of trips in Bin A, we compare them according to Bin B
- Whoever has more trips in Bin B goes earlier on the Student Ranking
- How do we order Bob and Kim?

	Bin A	Bin B
Bob	Trip 2, Trip 3	Trip 4, Trip 1
Kim	Trip 1, Trip 5	Trip 6

- **Bob has more trips in Bin B, so he goes higher than Kim on the Student Ranking**

## Creating the Student Ranking (summary)

- We start by trying to rank students according to Bin A. The person with more trips in Bin A goes higher on the Student Ranking.
- If we can’t rank two students based on Bin A because they have the same number of trips in Bin A, we try to rank them based on Bin B...etc.
- Let’s quickly see how we rank all 6 people in a group

## Assigning Students to Trips

- We proceed according to the Student Ranking, giving people trips from their most preferred bins
- We may swap people *within* bins to accommodate others

SR	Bin A	Bin B	Bin C
1 <sup>st</sup> : Jane	Trip 3, Trip 5	Trip 6	Trip 1
2 <sup>nd</sup> : Jeff	Trip 2	Trip 5, Trip 3	Trip 4, Trip 1
3 <sup>rd</sup> : Bob	Trip 2	Trip 3, Trip 4	Trip 6
4 <sup>th</sup> : Sarah	Trip 3	Trip 1	Trip 4, Trip 2, Trip 6
5 <sup>th</sup> : Alex	Trip 3	Trip 5	Trip 6, Trip 1
6 <sup>th</sup> : Rachel	Trip 3	Trip 2	Trip 6, Trip 1

## Making assignments: summary

We order students according to the Student Ranking

One at a time, according to the Student Ranking, we give each student a trip from the best bin possible; if there are multiple trips in your bin that the algorithm could give you, it picks one randomly from the bin

To accommodate others, we may swap you around to different trips *within* a bin

## Summary

- This experiment is 20 periods
- You will see your private values for the trips
- You submit choices (place the trips into bins)
  - We create the Student Ranking by placing students earlier who place the most trips in Bin A, breaking ties using Bin B, then Bin C... etc.
  - One at a time, according to the Student Ranking, we give each student a trip from the best bin possible (deciding randomly which trip in the bin to give you)
  - To accommodate others, we may swap you around to different trips *within* a bin

### Additional slides for the *Random treatment*

## Assigning Students to Trips

- We proceed according to the randomly generated Student Ranking, giving people trips from their most preferred bins
- We may swap people *within* bins to accommodate others

SR	Bin A	Bin B	Bin C
1 <sup>st</sup> : Jane	Trip 2	Trip 6	Trip 1
2 <sup>nd</sup> : Jeff	Trip 3, Trip 5	Trip 2	Trip 4, Trip 1
3 <sup>rd</sup> : Bob	Trip 3	Trip 1	Trip 2
4 <sup>th</sup> : Sarah	Trip 2	Trip 3, Trip 4	Trip 4, Trip 2, Trip 6
5 <sup>th</sup> : Alex	Trip 3, Trip 2, Trip 4	Trip 5	Trip 6, Trip 1
6 <sup>th</sup> : Rachel	Trip 3	Trip 2	Trip 6, Trip 1

## The Benefits and Costs of Bundling

The more trips you have in earlier bins, the greater your chance of appearing earlier in the Student Ranking

**“Bundling” multiple trips together has two BENEFITS**

1. Going earlier and possibly getting a trip from a more preferred bin
2. Allowing the algorithm to swap you around *within* your bin after you get a trip, which accommodates others

**However, “bundling” has the COST not necessarily getting your favorite trip within a bin; the algorithm**

1. chooses a random available trip from your bin, and may
2. swap you around *within* your bin after you get a trip to accommodate others

## Creating an Assignment: Overview

1. The algorithm **RANDOMLY** creates a “Student Ranking” of you and your group members; someone will be 1<sup>st</sup>, someone else 2<sup>nd</sup>, someone else 3<sup>rd</sup>...etc (You will not be told your ranking)
2. You make your choices (i.e. place the trips in bins)
3. According to the Student Ranking, the algorithm gives people their most preferred trips

## Making assignments: summary

We randomly order students to get a Student Ranking

One at a time, according to the Student Ranking, we give each student a trip from the best bin possible; if there are multiple trips in your bin that the algorithm could give you, it picks one randomly from the bin

To accommodate others, we may swap you around to different trips *within* a bin

## The Benefits and Costs of Bundling

## Summary

### “Bundling” multiple trips together has a BENEFIT of

- Allowing the algorithm to swap you around *within* your bin after you get a trip, which accommodates others

### However, “bundling” has the COST not necessarily getting your favorite trip within a bin; the algorithm

1. chooses a random available trip from your bin, and may
2. swap you around *within* your bin after you get a trip to accommodate others

- This experiment is 20 periods
- You will see your private values for the trips
  - We randomly order you and your group members to create a Student Ranking
  - You place the trips into bins
  - One at a time, according to the Student Ranking, we give each student a trip from the best bin possible (deciding randomly which trip in the bin to give you)
  - To accommodate others, we may swap you around to different trips *within* a bin

## References

- Abdulkadiroğlu, Atila, Sönmez, Tayfun, 1998. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica* 66 (3), 689–701.
- Abdulkadiroğlu, Atila, Che, Yeo-Koo, Yasuda, Yusuke, 2011. Resolving conflicting preferences in school choice: the “Boston” mechanism reconsidered. *Am. Econ. Rev.* 101, 399–410.
- Abdulkadiroğlu, Atila, Che, Yeo-Koo, Yasuda, Yusuke, 2015. Expanding “choice” in school choice. *Am. Econ. J. Microecon.* 7 (1), 1–42.
- Bergemann, Dirk, Morris, Stephen, 2005. Robust mechanism design. *Econometrica* 73 (6), 1771–1813.
- Bergemann, Dirk, Morris, Stephen, 2012. Ex post implementation. In: *Robust Mechanism Design: the Role of Private Information and Higher Order Beliefs*. World Scientific, pp. 97–152.
- Bogomolnaia, Anna, Moulin, Herve, 2001. A new solution to the random assignment problem. *J. Econ. Theory* 100 (2), 295–328.
- Budish, Eric, 2011. The combinatorial assignment problem: approximate competitive equilibrium from equal incomes. *J. Polit. Econ.* 119 (6), 1061–1103.
- Budish, Eric B., Kessler, Judd B., 2014. Changing the Course Allocation Mechanism at Wharton. Chicago Booth Research Paper (15-08).
- Calsamiglia, Caterina, Haeringer, Guillaume, Klijn, Flip, 2010. Constrained school choice: an experimental study. *Am. Econ. Rev.*, 1860–1874.
- Camerer, Colin F., 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press.
- Carroll, Gabriel, 2011. On Mechanisms Eliciting Ordinal Preferences. Technical report, Working Paper. MIT.
- Chen, Yan, Sönmez, Tayfun, 2002. Improving efficiency of on-campus housing: an experimental study. *Am. Econ. Rev.*, 1669–1686.
- Chen, Yan, Sönmez, Tayfun, 2004. An experimental study of house allocation mechanisms. *Econ. Lett.* 83 (1), 137–140.
- Chen, Yan, Sönmez, Tayfun, 2006. School choice: an experimental study. *J. Econ. Theory* 127 (1), 202–231.
- Coffman, Lucas C., Featherstone, Clayton R., Kessler, Judd B., 2017. Can social information affect what job you choose and keep? *Am. Econ. J. Appl. Econ.* 9 (1), 96–117.
- Edelman, Benjamin, Ostrovsky, Michael, Schwarz, Michael, 2007. Internet advertising and the generalized second-price auction: selling billions of dollars worth of keywords. *Am. Econ. Rev.* 97 (1), 242–259.
- Featherstone, Clayton R., 2014. Rank Efficiency: Investigating a Widespread Ordinal Welfare Criterion. Technical report, Working Paper.
- Featherstone, Clayton R., Niederle, Muriel, 2016. Boston versus deferred acceptance in an interim setting: an experimental investigation. *Games Econ. Behav.* 100, 353–375.
- Fischbacher, Urs, 2007. z-tree: Zurich toolbox for ready-made economic experiments. *Exp. Econ.* 10 (2), 171–178.
- Gale, David, Shapley, Lloyd Stowell, 1962. College admissions and the stability of marriage. *Am. Math. Mon.* 69 (1), 9–15.
- Holzman, Ron, Monderer, Dov, 2004. Characterization of ex post equilibrium in the vcg combinatorial auctions. *Games Econ. Behav.* 47 (1), 87–103.
- Hugh-Jones, David, Kurino, Morimitsu, Vanberg, Christoph, 2014. An experimental study on the incentives of the probabilistic serial mechanism. *Games Econ. Behav.* 87, 367–380.
- Hylland, Aanund, Zeckhauser, Richard, 1979. The efficient allocation of individuals to positions. *J. Polit. Econ.* 87 (2), 293–314. <https://doi.org/10.2307/1832088>. <http://www.jstor.org/stable/1832088>.
- Klijn, Flip, Pais, Joana, Vorsatz, Marc, 2013. Preference intensities and risk aversion in school choice: a laboratory experiment. *Exp. Econ.* 16 (1), 1–22.
- Li, Shengwu, 2017. Obviously strategy-proof mechanisms. *Am. Econ. Rev.* 107 (11), 3257–3287.
- Lien, Jaimie W., Zheng, Jie, Zhong, Xiaohan, 2015. Preference submission timing in school choice matching: testing fairness and efficiency in the laboratory. *Exp. Econ.*, 1–35.
- Liu, Qingmin, Pycia, Marek, 2016. Ordinal Efficiency, Fairness, and Incentives in Large Markets.
- Miralles, Antonio, 2008. School Choice: The Case for the Boston Mechanism.
- Pycia, Marek, Troyan, Peter, 2016. Obvious Dominance and Random Priority.
- Roth, Alvin E., 2002. The economist as engineer: game theory, experimentation, and computation as tools for design economics. *Econometrica* 70 (4), 1341–1378.
- Roth, Alvin E., Sönmez, Tayfun, Ünver, M. Utku, 2004. Kidney exchange. *Q. J. Econ.* 119 (2), 457–488.
- Shapley, L.S., Scarf, H., 1974. On cores and indivisibility. *J. Math. Econ.* 1, 23–28.
- Sönmez, Tayfun, 2013. Bidding for army career specialties: improving the rotc branching mechanism. *J. Polit. Econ.* 121 (1), 186–219.
- Sönmez, Tayfun, Switzer, Tobias B., 2013. Matching with (branch of choice) contracts at the United States military academy. *Econometrica* 81 (2), 451–488.
- Sönmez, Tayfun, Ünver, M. Utku, 2010a. Course bidding at business schools. *Int. Econ. Rev.* 51 (1), 99–123.
- Sönmez, Tayfun, Ünver, M. Utku, 2010b. House allocation with existing tenants: a characterization. *Games Econ. Behav.* 69 (2), 425–445.
- Sönmez, Tayfun, Ünver, Utku M., 2005. House allocation with existing tenants: an equivalence. *Games Econ. Behav.* 52 (1), 153–185.
- Thakral, Neil, 2015. Matching with Stochastic Arrival. Working paper. Harvard University.
- Troyan, Peter, 2012. Comparing school choice mechanisms by interim and ex-ante welfare. *Games Econ. Behav.* 75 (2), 936–947.
- Zhou, Lin, 1990. On a conjecture by gale about one-sided matching problem. *J. Econ. Theory* 52 (1), 123–135.