Journal of Mechanism and Institution Design

CONTENTS

Volume 3, Issue 1, December 2018
ISSN: 2399-844X (Print), 2399-8458 (Online)

Editorial board

Published by
The Society for the Promotion of Mechanism and Institution Design
Editorial office, Centre for Mechanism and Institution Design
University of York, Heslington, York YO10 5DD
United Kingdom
http://www.mechanism-design.org

ISSN: 2399-844X (Print), 2399-8458 (Online), DOI: 10.22574/jmid

The founding institutional members are

University of York, UK
Alpen-Adria-Universität Klagenfurt, Austria
Southwestern University of Economics and Finance, China.

Cover & Logo Artwork @ Jasmine Yang
LaTeX Editor & Journal Design @ Paul Schweinzer (using 'confproc')
LaTeX Editorial Assistant @ Theresa Marchetti

Printed in York and Klagenfurt by the Centre for Mechanism and Institution Design

A Letter from the Editor

With this third volume of *the Journal*, I want to share with you several items of interest concerning the *Journal of Mechanism and Institution Design* and its *Society for the Promotion of Mechanism and Institution Design*.

Firstly, *the Society* held its inaugural conference on 12th-13th May 2018. The conference was hosted by the prestigious Business School, Durham University, the third oldest university in England. Three distinguished economists were the keynote speakers of the conference: Vincent Crawford of the University of Oxford, David Martimort of Paris School of Economics, and Herve Moulin of the University of Glasgow. Besides the three keynote lectures, there were 48 invited parallel presentations and about 60 invited participants in total. (The conference was by invitation only.) It turned out to be a hugely successful meeting. We wish to thank local organisers Daniel Z. Li, Bibhas Saha, Ayse Yazici, Xiaogang Che, Simone Tonin, Anil Yildizparlak, and their colleagues and their school for their efforts and support in organising this conference.

Secondly, *the Society* has decided that its next conference will take place on 11th-13th June 2020 at the Department of Economics, Alpen-Adria-University of Klagenfurt, Klagenfurt, Austria. The University is situated a few hundred meters from the attractive Lake Wörthersee, is well-connected to Vienna, and only a short distance from both Slovenia and Italy across the Karawanken mountain range. We are very much looking forward to the event which will open for submissions in early 2019. You can find detailed conference information on the Journal website `http://www.mechanism-design.org/news.php`.

Thirdly, starting from January 2019, Yuan Ju will retire from his position as a co-editor, and Stephanie Lau, Alejandro Saporiti, Dolf Talman, and Akihisa Tamura from their positions as associate editors. We would like to express our heartfelt gratitude to them for their three-year excellent service, advice, and support which has been very important for the development and success of *the Journal*.

Fourthly, we are very pleased to announce that five outstanding colleagues have accepted our invitation to join our editorial board from January 2019. They are David Martimort of Paris School of Economics as a new co-editor, and Elizabeth Baldwin of the University of Oxford, Onur Kesten of Carnegie Mellon University, Scott Kominers of Harvard University, and Alex Teytelboym of the University of Oxford as new associate editors. We thank them for their willingness to serve the Journal and look forward to working with them closely to make *the Journal* a valuable outlet for our profession.

The first and ultimate goal of *the Journal* has been and will always be the same: To publish high quality papers in the field of mechanism and institution design and to provide the best possible services to the profession in the public interest. The Journal is completely free of charge, open-access to everyone, and makes no profit. It does not require submission, publication, or membership fees and thus goes beyond existing open-access journals. But the journal relies on donations and has received generous financial support. We hope that as time goes on, its mission will be shared by more and more people and institutions. We thank you for your support.

Zaifu Yang, York, 9th December, 2018

Journal of Mechanism and Institution Design
ISSN: 2399-844X (Print), 2399-8458 (Online)
DOI:10.22574/jmid.2018.12.001

THE UNCOVERED SET AND THE CORE: COX'S RESULT REVISITED

Anindya Bhattacharya

University of York, United Kingdom
anindya.bhattacharya@york.ac.uk

Victoria Brosi

Formerly of University of York, United Kingdom
vbrosi@gmx.de

Francesco Ciardiello

University of Sheffield, United Kingdom
f.ciardiello@shef.ac.uk

ABSTRACT

In this work first it is shown, in contradiction to the well-known claim in Cox (1987), that the uncovered set in a multidimensional spatial voting situation (under the usual regularity conditions) does not necessarily coincide with the core even when the core is singleton: in particular, the posited coincidence result, while true for an odd number of voters, may cease to be true when the number of voters is even. Second we provide a characterisation result for the case with an even number of voters: a singleton core is the uncovered set in this case if and only if the unique element in the core is the Condorcet winner.

We are greatly indebted to the Editor and two anonymous referees for several helpful comments and suggestions. We also thank Bhaskar Dutta, Gary Cox, John Duggan, Michel Le Breton and Elizabeth Maggie-Penn for their comments and help. This work was born when Brosi was receiving a scholarship from ESRC, UK and Ciardiello visited the Department of Economics and Related Studies, University of York. Brosi thanks the ESRC. Ciardiello thanks Regione Puglia for funding his International visiting scholarship. He thanks Andrea Genovese, head of the division "Operations Management and decision sciences" at Sheffield Management School, for funding recent short visiting period in York. He thanks the Department of Economics and Related Studies in York, especially Annette Johnson and Peter Simmons, for the kind hospitality always received in York. The errors and shortcomings remaining are ours.

Keywords: Spatial voting games, uncovered set, core.

JEL Classification Numbers: D71, C71.

1. INTRODUCTION

In this work first it is shown, in contradiction to the well-known claim in Cox (1987, 409) (repeated in a number of subsequent works), that the uncovered set in a multidimensional spatial voting situation does not necessarily coincide with the core even when the core is non-empty or even singleton: in particular, the posited coincidence result may cease to be true when the number of voters is even. Then we provide a characterisation result for the case with an even number of voters: a singleton core is the uncovered set in this case if and only if the unique element in the core is the Condorcet winner.

In our framework, the set of outcomes or policies under consideration is some compact and convex subset of some finite dimensional Euclidean space and any majority coalition of voters can enforce any outcome over another. For such an environment Cox (1987, 409) made the claim that if individual preferences satisfy a very innocuous symmetry condition then the uncovered set coincides with the core whenever the latter is non-empty. However, he worked with an odd number of voters *for expositional convenience* (Cox, 1987, 409). But actually his proof used the assumption that the cardinality of the voter set is odd in a non-trivial way. This claim has been repeated in subsequent literature. For example, Austen-Smith & Banks (2005, 274) stated in their much-used textbook that "the uncovered set coincides with the core when the latter is nonempty and singleton" (the definition of the uncovered set they use, in fact, gives a superset of the uncovered set with which we have worked here). A similar remark appears in the relatively recent but well-known paper by Penn (2009, 44).

However, in this paper we show that there is a voting situation for which the number of voters is even, the core is a singleton, but the uncovered set does not coincide with the core. Therefore, the message of this counter-example is that this claim (Cox, 1987), which is a powerful result for majority rule voting situations with an odd number of voters, should be invoked with some degree of caution. The strength of this result in the case with an odd number of voters (for which it is true) may lead to its somewhat careless generalised

use. While we show this possibility of non-coincidence of the core and the uncovered set for spatial voting, for majority voting situations with a finite set of outcomes, this possibility was noted by Bordes (1983) several years ago. Quite naturally, the question then arises whether an analogue of Cox (1987)'s result is true when the number of voters is even. Addressing this issue we provide the following characterization result: a singleton core is the uncovered set in a majority spatial voting situation (with the usual regularity conditions) with an even number of voters if and only if the unique element in the core is the Condorcet winner of that situation.

The next section gives the preliminary definitions and notation. Section 3 gives the results and some discussions about the results.

2. PRELIMINARY DEFINITIONS AND NOTATION

Let $Z \subseteq \mathbb{R}^k$ be a compact convex subset of some finite (k-)dimensional Euclidean space. This set, Z, is identified to be the feasible set of policies or outcomes on which a voter votes. Let $N = \{1, 2, \ldots, n\}$ be the finite set of players or voters. Suppose that the preferences of a player i on Z is represented by a real-valued continuous and strictly concave pay-off function $u_i \in C^0(Z, \mathbb{R})$. The spatial voting situation we consider below is obtained by introducing the method of majority rule voting.

Definition 2.1 (Domination by Majority Rule)**.** *Given $x, y \in Z$, the policy x beats (or dominates) policy y via coalition $S \subseteq N$, if $|S| > |N|/2$ and $u_i(x) > u_i(y)$ for each $i \in S$. We denote this as $x \succ_S y$. If there exists a majority coalition S via which x dominates y, we denote that as $x \succ y$.*

The collection $G = (Z, N, (u_i)_{i \in N})$ is a spatial voting situation with majority rule. For any $x \in Z$ and $i \in N$, by $D^i(x)$ we denote the set $\{y \in Z : u_i(y) > u_i(x)\}$. Further, $D(x) = \{y \in Z : y \succ x\}$. For any set $A \subseteq Z$, by $cl(A)$ we denote the closure of A. Also, for any two points $x, y \in Z$, by $\rho(x, y)$ we denote the (Euclidean) distance between these two points. Recall the two well-known solution concepts for such situations that we shall discuss: the core and the uncovered set.

Definition 2.2 (The Core of a Voting Situation)**.** *The core of such a voting situation is the subset $K = \{y \in Z : \nexists z \in Z \text{ such that } z \succ y\}$.*

Recall that a point $x \in Z$ is said to be the Condorcet winner of the voting situation if for any other outcome $y \neq x$, $x \succ y$. Recall that if a voting situation admits a Condorcet winner, then it is the unique element in the core.

Definition 2.3 (The (Gillies) Uncovered Set)**.** *Let $x, y \in Z$. We say that x covers y, denoted as $y \prec_c x$ if the following hold:*

$$x \succ y;$$
$$z \in Z,\ z \succ x \Longrightarrow z \succ y.$$

The uncovered set is given by $UC = \{y \in Z : \nexists z \text{ such that } y \prec_c z\}$

As this work is primarily motivated by Cox (1987) we are using the definition he has used. Although the notion of covering (and that of uncovered sets) was introduced explicitly first by Miller, the solution defined here is the set of the maximal elements of the Gillies' covering subrelation (following Gillies, 1959) rather than Miller's subrelation (for clarification, we refer to Bordes et al. 1992; see also Miller 2015). At the end we remark on the analogue of our non-coincidence result for the (Miller) uncovered set. For any $y \in Z$, by $C(y)$ we denote the set of elements in Z, which cover y.

Next recall that the preferences are said to be Euclidean or circular if for every voter $i \in N$, there exists $\bar{x}_i \in Z$ such that for any policy $x \in Z$, $u_i(x) = -(\rho(x, \bar{x}_i))^2$. Cox (1987, 416) introduces a notion of *limited* asymmetry of preferences. For completeness, we reproduce the definition here.

Definition 2.4 (Preferences that are limited in asymmetry)**.** *Preferences are said to be limited in asymmetry by α if for every line L intersecting Z, every $i \in N$, and every $r \in \mathbb{R}$,*

$$V_L^i(r) \neq \emptyset \Longrightarrow f(L, i.r) \leq \alpha$$

where $V_L^i(r) = \{x \in L : u_i(x) = r\}$; $f(L, i, r) = \frac{max_{x \in V_L^i(r)} \rho(x, b_L^i)}{min_{x \in V_L^i(r)} \rho(x, b_L^i)}$; the induced ideal point of $i \in N$ on the line L being denoted by b_L^i.

We merely note here that Euclidean preferences obviously satisfy this condition with $\alpha = 1$.

3. THE RESULTS AND DISCUSSIONS

Cox's result is as follows.

Proposition 3.1. *Take a voting situation G. Suppose there exist a finite α such that the condition of limited asymmetry of preferences by α holds for every voter. Now suppose $K \neq \emptyset$. Then $K = UC$.*

As we mentioned in the introduction, one primary motivation behind this work is that this result has been repeated in a number of subsequent works but not with, in our view, sufficient care. However, we find the following.

Proposition 3.2. *There is a voting situation G for which $|N|$ is even, the core K is singleton, the preference of each of the voters is Euclidean and the uncovered set does not coincide with the core.*

To prove this proposition we shall use an intermediate result. First recall the definition of a von-Neumann-Morgenstern stable set for the voting situations we consider here.

Definition 3.1 (von-Neumann-Morgenstern Stable Sets)**.** *A set $V \subseteq Z$ is a (von-Neumann-Morgenstern) stable set for G if it satisfies*

- *(internal stability:) there do not exist $x, y \in V$ such that $x \succ y$;*
- *(external stability) if $x \in Z \setminus V$ it must be the case that there exists $y \in V$ such that $y \succ x$.*

The following Proposition 3.3 is useful to prove Proposition 3.2. Although variants of this result are well-known (see, e.g., McKelvey (1986)), for completeness we provide a short proof, within our framework, below.

Proposition 3.3. *If a stable set V exists then $K \subseteq V \subseteq UC$.*

Proof. If $K \not\subseteq V$ then that violates the external stability of V.

Let V be a stable set and take, if possible, and $x \in V \setminus UC$. That is, there exists $y \in Z$ such that $x \prec_c y$. This implies that $y \succ x$. Since, $y \notin V$ (otherwise, the internal stability of V is violated), by external stability of V, there exists $z \in V$ such that $z \succ y$. But, then, by the definition of the covering relation, $z \succ x$ which again violates the internal stability of V. □

Proof of the Proposition 3.2. Below we give an example of a situation where the core is singleton, a stable set exists and the stable set does not coincide with the core. Then, by Proposition 3.3 we are done. Let $N = \{1,2,3,4\}$. The set of outcomes, $Z = \{x \in \mathbb{R}^2 | x_1 \in [-1,1]; x_2 \in [-1,1]\}$. Each player i has an ideal point $\bar{x}_i$ whose coordinates are given as follows. The point $\bar{x}_1$ (labelled by A')$= (-1,-1)$; $\bar{x}_2$ (labelled by B')$= (1,-1)$; $\bar{x}_3$ (labelled by C')$= (1,1)$ and $\bar{x}_4$ (labelled by D')$= (-1,1)$ (please see Figure 1 below). The players' preferences are Euclidean, i.e., for any $i \in N$, and $x \in Z$, $u_i(x) = -(\rho(x,\bar{x}_i))^2$. We show below that the core of this situation is the singleton set containing the point $(0,0)$ (labelled as point O) while the set $V = \{x \in Z | x_1 = 0 \text{ or } x_2 = 0\}$ is a stable set. For convenience later call the set $\{x \in Z | x_1 = 0\}$ as V_1 and the set $\{x \in Z | x_2 = 0\}$ as V_2.

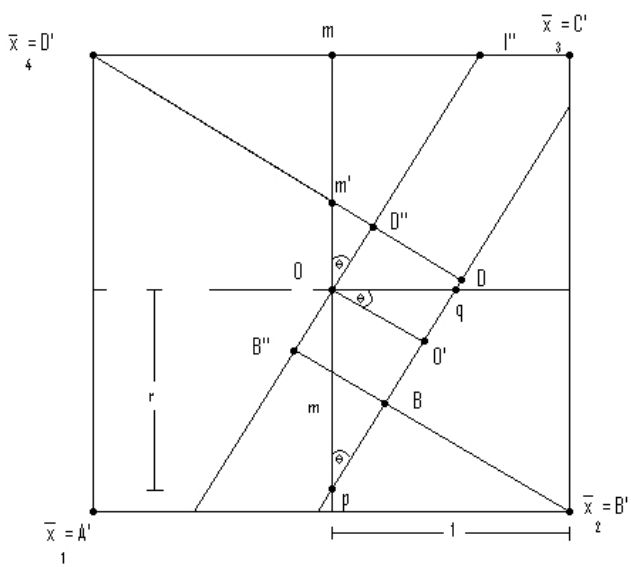

Figure 1: The voting situation for which the singleton core is not a stable set.

We take the following steps.

Step 1 : Notice that the point $O = (0,0)$ satisfies the Plott condition (Austen-Smith & Banks, 1999, 142) and so, it is in the core of this voting situation (see Schofield, 2008, 90), too.

Step 2 : Next we show that no point other than O is in the core. We start with the subset $\Delta_1 = \{x \in Z \setminus \{O\} | x_1 \leq 0, x_2 < 0, x_1 \geq x_2\}$. Note that this is a triangle (without the point O. Choose a point $x \in \Delta_1$ such that $x_1 < 0, x_2 < 0$ and $x_1 > x_2$. Draw a line of slope 1, L_x, passing through this x and let the line intersect the line V_1 at the point y. It is obvious that y dominates x via coalition $\{2,3,4\}$. Next, take a point $x \in \Delta_1$ such that $x_1 = x_2$. Then it is obvious that $(0,0)$ dominates such a point via $\{2,3,4\}$. Finally, choose a point $x \in \Delta_1$ such that $x_1 = 0$ and $x_2 < 0$, i.e., the point is in V_1. Again, draw a line of slope 1, L_x, passing through x and let that intersect the line given by $\{x \in Z | x_1 = -x_2\}$ at a point y. Again, it is obvious that y dominates x via coalition $\{2,3,4\}$.
Thus, no point of Δ_1 is in the core.
Note that $Z \setminus \{(0,0)\}$ is the union of 8 such triangles like Δ_1. Therefore, using the symmetry between these triangles we can show that no point other than O is in the core.

Next we show that V is a stable set.

Step 3 (External stability of V): From Step 2 itself we see that for any $x \in Z \setminus V$, there exists an $y \in V$ such that $y \succ x$.

Step 4 (Internal stability of V): It is obvious that a point in V_1 cannot be dominated by another point in V_1 and a point in V_2 cannot be dominated by another point in V_2. Next we show that a point in V_1 cannot dominate, nor can be dominated by a point in V_2. Let p be a point in V_1 such that $p_2 < 0$. Let q be a point in V_2 such that $q_1 > 0$. Let $\overline{Op}$, the length of the line segment Op, be $0 < r \leq 1$ and let the angle $\widehat{Opq}$ be θ. (Please refer to Figure 1.) Let, without loss of generality, $0 < \theta \leq \pi/4$. Call the line passing through p and q, L. Let the perpendiculars from A', B', C' and D' on L (or, to put more rigorously, the orthogonal projections of these points on L) be denoted respectively by A, B, C and D. Note that if $\theta = \pi/4$, then the points B and D coincide and then it is obvious that, neither p can dominate q, nor q can dominate p. Now suppose

$0 < \theta < \pi/4$. It is easy to see that $\overline{qD} < \overline{pD}$. Since $\overline{Cq} < \overline{Cp}$, p cannot dominate q. Next we show that the line segment $\overline{B'p} \leq \overline{B'q}$. Note that this is true if and only if $(1+p_2) \leq (1-q_1)$. But this follows from the fact that $\theta < \pi/4$. Since $\overline{A'p} \leq \overline{A'q}$, (obviously) q cannot dominate p. From this, using the symmetry of this example, we can show that no point in V_1 can dominate another point in V_2 and vice versa. □

One point of curiosity is to identify the uncovered set in the example we used in proving Proposition 3.2. [1] Indeed, we find that for that example, the uncovered set is a strict superset of the stable set we identified, a feature which might be somewhat interesting. We summarize the finding as an additional result below.

Result 3.1 Consider the set U specified as:

$$\begin{gathered}(\{z \in Z \,|\, \rho(A',z) \leq \rho(A',O)\} \cap \{z \in Z \,|\, \rho(B',z) \leq \rho(B',O)\}) \cup \\ (\{z \in Z \,|\, \rho(B',z) \leq \rho(B',O)\} \cap \{z \in Z \,|\, \rho(C',z) \leq \rho(C',O)\}) \cup \\ (\{z \in Z \,|\, \rho(C',z) \leq \rho(C',O)\} \cap \{z \in Z \,|\, \rho(D',z) \leq \rho(D',O)\}) \cup \\ (\{z \in Z \,|\, \rho(D',z) \leq \rho(D',O)\} \cap \{z \in Z \,|\, \rho(A',z) \leq \rho(A',O)\}).\end{gathered}$$

The set U (shown as the shaded area in the Figure 2 below) is the uncovered set in the example used in Proposition 3.2.

The proof of Result 3.1 involves repeated (and somewhat tedious) use of similar arguments from elementary Euclidean geometry. We provide a sketch proof below.

[1] We are indebted to one of the referees for inducing us to explore this issue.

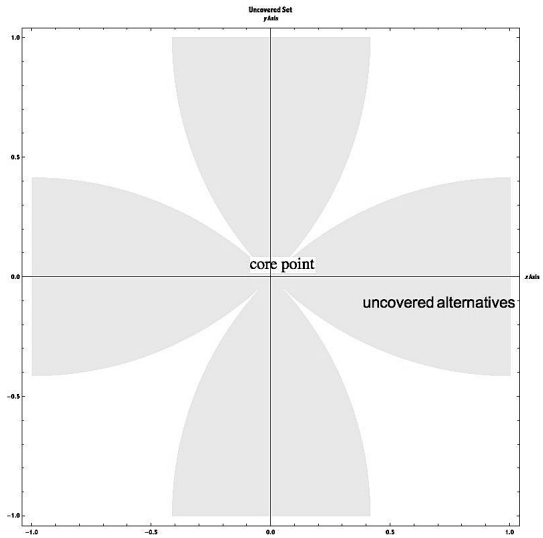

Figure 2: The uncovered set for the example used in Proposition 3.2.

Proof of Result 3.1. We start with defining the following subsets of Z :

$$\begin{aligned}
\Delta_1 &= \{x \in Z \,|\, x_1 < 0, x_2 < 0, x_1 > x_2\},\\
\Delta_2 &= \{x \in Z \,|\, x_1 < 0, x_2 < 0, x_1 < x_2\},\\
\Delta_3 &= \{x \in Z \,|\, x_1 > 0, x_2 > 0, x_1 > x_2\},\\
\Delta_4 &= \{x \in Z \,|\, x_1 > 0, x_2 > 0, x_1 < x_2\},\\
\Delta_5 &= \{x \in Z \,|\, x_1 > 0, x_2 < 0, x_1 > -x_2\},\\
\Delta_6 &= \{x \in Z \,|\, x_1 > 0, x_2 < 0, x_1 < -x_2\},\\
\Delta_7 &= \{x \in Z \,|\, x_1 < 0, x_2 > 0, -x_1 < x_2\},\\
\Delta_8 &= \{x \in Z \,|\, x_1 < 0, x_2 > 0, -x_1 > x_2\}.
\end{aligned}$$

Then the proof proceeds along the following steps.

Step 1 First note the rather obvious fact that no point $x \in Z$ is dominated by another point y, via the grand coalition $\{1,2,3,4\}$. For completeness we give a brief proof. Consider, without loss of generality, a point x in Δ_1. Again, without loss of generality, consider a point y in the "north-east" of x. Then $u_1(x) \geq u_1(y)$. By using similar arguments, and by the symmetry of the voting situation in this example, this fact can be verified.
Take a point $X \in \Delta_1 \cap \{y \in Z : \rho(y,B') < \rho(O,B')\}$ (refer to Figure 3

below as well). Below, for $y \in Z$ and a positive real number r, by $L(y,r)$ we denote the set $\{z \in Z : \rho(y,z) < r\}$.

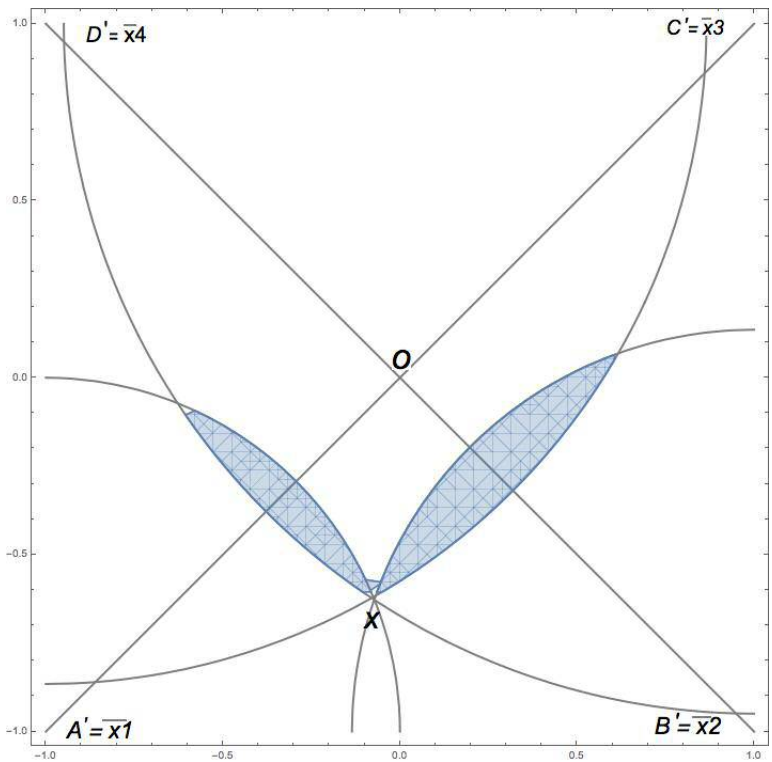

Figure 3: The set of points dominating the point X.

Step 2 Step 1 implies that for any $x, y \in Z$, if y dominates x, then that must be via a coalition of cardinality 3. Now take the point X without loss of generality. Note that if some $y \succ_S X$ via some coalition S, then $\{1,2\} \not\subset S$. To see this, assume the contrary. Then, there exists $y \in L(A', \rho(A',X) \cap L(B', \rho(B',X)$ which dominates X via some coalition S which contains voters 1 and 2. Let, without loss of generality, the third member of S be 3. Consider the angles $A'C'X = \phi$ and $A'C'y = \phi'$. It is easy to see that $\pi/4 > \phi' > \phi$. Then, from the fact that $\cos\phi' < \cos\phi$, it is straightforward that $\rho(C',y) > \rho(C',X)$. But this leads to a contradiction.

Step 3 It is easy to see that, then, $D(X) = (D^1(X) \cap D^3(X)) \cup (D^2(X) \cap D^4(X))$ (this is depicted as the shaded area, without the boundaries, in Firgure 3).

Step 4 Now consider, if possible, $z \in Z$, such that z can cover X. Suppose $z \in D^2(X) \cap D^4(X) \cap \Delta_1$. Then $D(z)$ must contain a point in $\Delta_1 \setminus (D^1(X) \cap D^3(X))$. Suppose $z \in D^2(X) \cap D^4(X) \cap \Delta_6$ such that $\rho(A', z) < \rho(A', O)$. Then, again, $D(z)$ must contain a point in $\Delta_1 \setminus (D^1(X) \cap D^3(X))$. Suppose $z \in D^2(X) \cap D^4(X) \cap \Delta_6$ such that $\rho(A', z) > \rho(A', O)$. Then, again, $D(z)$ must contain O, the unique point in the core, but $O \notin D(X)$. By replicating the arguments above in this fashion, we can verify that X cannot be covered.

Step 5 Steps like the above can be replicated for every point in the set U, which would verify that U is indeed the uncovered set. Note that a point of Z which is neither in the core and nor in U is covered by O, the single point in the core. □

Given Proposition 3.2 (and Result 3.1), the immediate question which obviously arises is what might be the conditions under which an analogue of Cox's result–coincidence of the core and the uncovered set–holds when the number of voters is even? We find the following.[2]

Proposition 3.4. *Consider a voting situation G for which $|N|$ is even and for which the core $K = \{x_0\}$ is singleton. Then K is the uncovered set if and only if x_0 is the Condorcet winner for the voting situation.*

The following elementary lemma would be useful for proving this proposition.

Lemma 3.1. *Suppose $x_0 \in Z$ is a point in the core and that it does not dominate a point $y \in Z; y \neq x_0$. Then there exists a coalition T of exactly $n/2$ voters such that for each $i \in T$, $u_i(x_0) > u_i(y)$.*

Proof. Suppose not. Then there exists a coalition S containing at least $n/2+1$ voters such that for every $i \in S$, $u_i(x_0) \leq u_i(y)$. Consider a point $z \neq x_0 \neq y$ such that z is a convex combination of x_0 and y. But then, since u_i is strictly

[2] We are especially indebted to the Editor for encouraging us to explore a result like Proposition 3.4.

concave for each $i \in N$, for each $i \in S$, $u_i(z) > u_i(x_0)$ which contradicts the supposition that x_0 is in the core. The proof is completed by noting the fact that if x_0 is strictly preferred to y by more than $n/2$ voters, then x_0 dominates y which, once again, leads to a contradiction. □

Proof of Proposition 3.4. The proof of the "if" part is obvious. The proof of the "only if" part proceeds along the following steps.

Step 1 Recall that the covering relation, $\prec_c$, is transitive: i.e., $z \prec_c y$ and $y \prec_c x$ implies that $z \prec_c x$ (see, e.g., Bordes et al. (1992)).

Step 2 Now pick, if possible, $y \in Z \setminus K$, such that x_0 does not dominate y. Then, setting $y^0 = y$, construct an infinite sequence of sets $(G^q)_{q=1}^{\infty}$ in the following manner:
$G^1 = C(y^0)$;[3]
then choose some member $y^1 \in G^1$, and
$G^2 = C(y^1)$;
and proceeding in this manner, given G^k and choosing some $y^k \in G^k$,
$G^{k+1} = C(y^k)$.
By the transitivity of $\prec_c$, for each k, $G^k \subset G^{k-1}$. Observe also, that for each k, $G^k \subseteq D(y^{k-1})$. Note that by the construction of the sequence, if for any k, $x_0 \in G^k$, then, by the transitivity of $\prec_c$, $x_0 \succ y$ and we reach an immediate contradiction. Therefore, for every k, $x_0 \notin G^k$. Further, by the same reasoning, by the construction of the sequence $(G^q)_{q=1}^{\infty}$,

$$\cap_{j=1}^{\infty} G^j = \emptyset.$$

Step 3 Then, further consider the closure of each of these sets in the sequence. Then the sequence of sets $(cl(G^j))_{j=1}^{\infty}$ is a sequence of non-empty compact subsets of Z such that for every j,

$$cl(G^{j+1}) \subseteq cl(G^j).$$

Then, by the Cantor intersection lemma (see, if necessary, e.g. (Rudin, 1976, 38)),

$$\cap_{j=1}^{\infty} cl(G^j) \neq \emptyset.$$

[3] Recall from Section 2 that for any $z \in Z$, by $C(z)$ we denote the set of points which cover z.

Therefore, there exists an $x \in Z$, such that x is in the closure of every G^j. Recall, however, from Step 2 above that $\cap_{j=1}^{\infty} G^j = \emptyset$. Therefore, for every $j \geq k$ (where k is some positive integer) there exists a sequence (y_j^r), with each element from the sequence in G^j, which converges to x.

Step 4 Pick such a G^j, i.e., for which $j \geq k$. By the definition of G^j, for every element $z \in G^j$, $D(z) \subset D(y^{j-1})$. Therefore, by Lemma 7.2 in Austen-Smith & Banks (2005, 272), $D(x) \subseteq D(y^{j-1})$.[4] Therefore,

$$D(x) \subseteq \cap_{j=1}^{\infty} D(y^{j-1}).$$

Note that by the construction of the sequence in Step 2 above,

$$\cap_{j=1}^{\infty} D(y^{j-1}) = \emptyset.$$

Therefore, $D(x) = \emptyset$, i.e., $x = x_0$.

Step 5 Therefore, $x_0 \in cl(G^1)$. Since $G^1 \subset D(y^0)$, $x_0 \in cl(D(y^0))$. If $x_0 \in D(y^0)$ (i.e., $D(y)$), then again we get an immediate contradiction. So, suppose otherwise.

Step 6 Then $x_0 \in cl(D(y)) \setminus D(y)$. Therefore, there exists a sequence (y^r), with each element from the sequence in $D(y)$, which converges to x_0. Further, since there are only finitely many coalitions, there exists a fixed majority coalition S and a sequence $(z^l)_{l=1}^{\infty}$ such that for each element z of the sequence, $z \in D(y)$, z dominates y via the coalition S, and the sequence (z^l) converges to x_0. Since x_0 does not dominate y via S, there exists a voter $i \in S$ such that for each z in the sequence, $u_i(z) > u_i(y) \geq u_i(x_0)$. Moreover, since z^l converges to x_0, by continuity of u_i, $u_i(y) = u_i(x_0)$.

Step 7 Recall that by Lemma 3.1, there exists a coalition T of exactly $n/2$ voters such that for each $i \in T$, $u_i(x_0) > u_i(y)$. Moreover, conversely, for each voter j in the $n/2$-voter coalition $N \setminus T$, $u_j(y) \geq u_j(x_0)$. Then, by strict concavity of the pay-off functions, for each $w \in Z$ which is a convex combination of x_0 and y, (and different from either x_0 or y) $u_j(w) > u_j(x_0)$ for every $j \in N \setminus T$. But then, since x_0 is in the core,

[4] Recall that the lemma, in the present notation-style, states: let (z^j) be an infinite sequence converging to $x \in Z$ such that for every j, $D(z^j) \subseteq D(y)$ for some $y \in Z$. Then $D(x) \subseteq D(y)$.

for each $j \in T$, $u_j(x_0) > u_j(w)$ for each such w on the linear segment x_0y (otherwise x_0 gets dominated). In particular, none of the w's on the segment x_0y is dominated by x_0. Pick one such w. But then, replicating the argument with respect to y above, given in Steps 2 to 6, there must exist some voter $k \in N$ for whom $u_k(x_0) = u_k(w)$. But this contradicts the fact that for every $i \in N$, either $u_i(x_0) > u_i(w)$ or $u_i(x_0) < u_i(w)$.

□

4. CONCLUDING REMARKS

Here, at the end, we provide a few remarks related to our results. First, in view of Proposition 3.4, notice that in the example used in Proposition 3.2, the single point in the core is not the Condorcet winner. Next, the fact that this non-coincidence is still true with, e.g., the Miller definition of the uncovered set, follows in a straightforward manner from Proposition 30 in Duggan (2013). And finally, there remain several open questions: e.g., how far a result like Proposition 3.4 can be generalised, under what primitive conditions a singleton core can never be a Condorcet winner, etc. These are topics of ongoing and further research.

References

Austen-Smith, D., & Banks, J. S. (1999). Positive Political Theory I. University of Michigan Press.

Austen-Smith, D., & Banks, J. S. (2005). Positive Political Theory II. University of Michigan Press.

Bordes, G. (1983). On the possibility of reasonable consistent majoritarian choice: Some positive results. Journal of Economic Theory, 31(1), 122–132.

Bordes, G., Le Breton, M., & Salles, M. (1992). Gillies and Miller's subrelations of a relation over an infinite set of alternatives: general results and applications to voting games. Mathematics of Operations Research, 17(3), 509–518.

Cox, G. W. (1987). The uncovered set and the core. American Journal of Political Science, 31(2), 408–422.

Duggan, J. (2013). Uncovered sets. Social Choice and Welfare, 41(3), 489–535.

Gillies, D. B. (1959). Solutions to general non-zero-sum games. In A. Tucker & R. Luce (Eds.), Contributions to the theory of games IV (pp. 47–85). Princeton University Press.

McKelvey, R. D. (1986). Covering, dominance, and institution-free properties of social choice. American Journal of Political Science, 30(2), 283–314.

Miller, N. R. (2015). The spatial model of social choice and voting. In J. Heckelman & N. Miller (Eds.), Handbook of Social Choice and Voting (pp. 163–181). Edward Elgar.

Penn, E. M. (2009). A model of farsighted voting. American Journal of Political Science, 53(1), 36–54.

Rudin, W. (1976). Principles of Mathematical Analysis. McGraw-Hill.

Schofield, N. (2008). The Spatial Model of Politics. Routledge.

Journal of Mechanism and Institution Design
ISSN: 2399-844X (Print), 2399-8458 (Online)
DOI:10.22574/jmid.2018.12.002

STRATEGY-PROOFNESS OF STOCHASTIC ASSIGNMENT MECHANISMS

André Schmelzer

Max Planck Institute for Research on Collective Goods, Germany
schmelzer.andre@gmail.com

ABSTRACT

This paper seeks to identify which algorithm to employ in a situation where goods are distributed to individuals without using money, while treating everyone equally and respecting each individual's preferences. I compare two stochastic assignment mechanisms: Random serial dictatorship (RSD) and top trading cycles with random endowments (TTC). In standard theory, both algorithms are strategy-proof and yield the optimal result. In the experiment, RSD outperforms TTC. This can be attributed to a more dominant strategy play under RSD. Generally, subjects with extremely high and low levels of contingent reasoning play their dominant strategies. These results suggest that one optimal algorithm may outperform another one if individuals are boundedly rational.

Keywords: Market design, mechanism design, randomization.

JEL Classification Numbers: C78, C92, D47.

1. INTRODUCTION

Which algorithm should be employed in a situation where goods are distributed to individuals without using money? I compare algorithms

I thank Dorothea Kübler, Christoph Engel, Rustamdjan Hakimov, Yoan Hermstrüwer, Svenja Hippel, Johannes Hemker, Tim Winke, and seminar participants at the Max Planck Institute for Research on Collective Goods Bonn and Humboldt University of Berlin for helpful discussions and suggestions. Financial support from the Max Planck Society for the Advancement of Science is gratefully acknowledged. The author is an economist at the German Federal Ministry of Finance. The views expressed in this paper are those of the author and do not necessarily reflect those of the Ministry of Finance.

that perform such an assignment, while treating everyone equally and respecting each individual's preferences over the goods to be distributed. In principle, the class of stochastic assignment mechanisms in this paper is regarded as fair because chance treats everyone equally. This fairness property along with respecting individuals' preferences is key for applications. However, stochastic assignment mechanisms may not yield the optimal result if individuals do not truthfully reveal their preferences. What may be worse is that individuals who do not understand the dominant strategy could be disadvantaged. Then, the social planner faces a challenging choice between optimal mechanisms.

In this paper, I study mechanisms for this stochastic assignment of n indivisible objects to n agents without existing endowments or monetary transfers. Two prominent one-sided matching mechanisms are compared in the laboratory: *Random serial dictatorship* (RSD) and Gale's *top trading cycles algorithm* (Shapley & Scarf, 1974) with random endowments (TTC). In RSD, agents are randomly ordered and then objects are assigned in this order according to their preference. In TTC, initially objects are randomly allocated to agents. Then, mutually beneficial exchanges are performed based on agents' preferences. Both mechanisms are strategy-proof and ex-post Pareto-efficient (Abdulkadiroğlu & Sönmez, 1998). In standard theory, they are equivalent.

In practice, it is less clear whether this equivalence holds with boundedly rational agents. Moraga & Rapoport (2014) proposed to implement TTC and RSD for refugee resettlement. Accordingly, indivisible goods (residence permits) are assigned to agents (refugees) without endowments (pre-existing right to enter a country). Other applications include time slots to users of a common machine, night shifts to doctors, or public housing to tenants. In such real life problems, understandability is key since the efficiency of the outcome depends on the agents' ability to comprehend the dominant strategy.[1] The social planners' choice between theoretically identical mechanisms matters if the dominant strategy is easier to recognize under one mechanism than under the other. Up to now, there is no evidence which allows a comparison of TTC and RSD without endowments.

Assuming bounded rationality, the theory of *obvious strategy-proofness* (OSP) (Li, 2017) predicts differences in dominant strategy play between RSD and TTC. A mechanism is obviously strategy-proof if recognizing the dominant

[1] In the field, failures to recognize dominant strategies are partly found to be strategically motivated (Rees-Jones, 2017) and persist even if information about the dominant strategy is provided (Hassidim et al., 2016).

strategy does not require contingent reasoning about the hypothetical actions of others. The *sequential* version of RSD (OSP-RSD) is obviously strategy-proof because agents are taking turns and choose an object from a set, one after another. They play the dominant strategy by picking the highest prize. When making their choice, they do not need to reason contingently about the hypothetical choices of the others to play the dominant strategy. In contrast, TTC and the *simultaneous* version of RSD (SP-RSD) are not obviously strategy-proof because playing the dominant strategy requires contingent reasoning. To recognize the dominant strategy when submitting their rank order list to the mechanisms, the agents need to think about the hypothetical lists of the other agents contingent on their own list. Thus, OSP predicts for boundedly rational agents who do not perfectly reason contingently that the frequency of dominant strategy play is larger in OSP-RSD compared to TTC and SP-RSD.

In this paper, I test the performance of TTC and RSD in the laboratory. I investigate whether dominant strategy play differs between mechanisms and whether the Pareto-efficient outcome is attained. The experiment is designed to compare TTC with both versions of RSD; with the simultaneous SP-RSD and with the sequential, obviously strategy-proof OSP-RSD.

The results clearly show that RSD outperforms TTC. The Pareto-efficient welfare level is attained more often in SP-RSD than in TTC. This can be attributed to more dominant strategy play in the RSD mechanisms compared to TTC. This contrasts with the predictions of standard game theory. As predicted by OSP, there is more dominant strategy play in OSP-RSD than in TTC. However, OSP fails to explain why there is more dominant strategy play in SP-RSD compared to TTC and no differences between OSP-RSD and SP-RSD.

These findings complement the matching literature on the house allocation problem with existing endowments, e.g., squatting rights. Abdulkadiroğlu & Sönmez (1999) show that in this case TTC is Pareto-efficient, but RSD is not, because it is not individually rational for every agent to participate in RSD in the first place. Laboratory experiments under incomplete (Y. Chen & Sönmez, 2002) and complete information (Y. Chen & Sönmez, 2004) find that, in line with this theory, RSD is less efficient than TTC. This problem is also known as the housing market with old and new tenants. In contrast, I study the case with *new* tenants only. This special case is relevant whenever agents do not have prior claims for an object like in assigning slots, licenses, and permits.

The closest related paper is Li (2017). He formulates the theory of OSP,

taking the limited ability for contingent reasoning into account, and tests it experimentally.[2] He finds that dominant strategy play and efficiency are larger in sequential RSD than in simultaneous RSD. In addition, my paper directly compares TTC with both versions of RSD. With this comparison, I complement the literature indicating that individuals fail to always play dominant strategies in TTC (e.g., Guillen & Hakimov, 2016) and in RSD (Olson & Porter, 1994).

In the experiment, I introduce a new method to compare sequential and simultaneous mechanisms. My aim is to make decisions more comparable between simultaneously submitting a list of preferences and sequentially selecting an object from a set. I argue that comparing dominant strategy play between a list of 4 preferences with a single choice of an object is not fair because subjects are more likely to make an error when submitting a list. I assume that ranking 4 objects in a list involves making a choice about each individual rank of an object, e.g., by pairwise comparisons. I use the strategy method (Selten, 1967) to implement the sequential OSP-RSD: Subjects choose from 4 different sets of objects without knowing their position in the random sequence. Assuming that subjects evaluate each choice from a set of objects separately (Rabin & Weizsäcker, 2009), the procedure remains OSP. As a result, a preference list containing 4 objects is comparable with 4 choices from sets of objects. It turns out that this method yields different results compared to earlier works. I replicate the finding of Li (2017) that OSP-RSD yields more dominant strategy play than SP-RSD, when I use one single choice from one set of objects using the strategy method data. However, the error rate increases if I use all 4 choices from 4 sets, resulting in the difference between the sequential OSP-RSD and the simultaneous SP-RSD vanishing.

To shed light on the behavioral mechanism behind dominant strategy play, I identify extreme forms of contingent reasoning based on two additional games. At one extreme, I identify subjects who are perfectly able to reason hypothetically about the actions of others by guessing 0 in the 2-person beauty contest game (Grosskopf & Nagel, 2008). At the other extreme, I identify subjects would who rather refrain from hypothetical thinking about the actions

[2] Theory on OSP is evolving. For instance, Pycia & Troyan (2018) introduce a refinement by showing that the class of sequential dictatorships is *strong* OSP. Troyan (2016) characterizes an OSP implementation of TTC beyond the housing allocation problem. Ashlagi & Gonczarowski (2016) show that the Gale-Shapley deferred acceptance algorithm is not OSP. L. Zhang & Levin (2017) provide an axiomatization of the failure to reason state-by-state. Related, Glazer & Rubinstein (1996) find that the implementation of a social choice function via a normal-form game can be more obvious under certain conditions.

of others by requesting 20 in the 11-20 money request game (Arad & Rubinstein, 2012). The upside of this second game is that requesting 20 does not require contingent reasoning, but is at the same time a sensible answer when not knowing what the other person requests. Both extreme types play their dominant strategies. In this way, my work relates to the growing literature on contingent reasoning in other settings such as school choice (J. Zhang, 2016), takeover games (Charness & Levin, 2009), financial markets (Ngangoue & Weizsäcker, 2015), voting (Esponda & Vespa, 2014), and in explaining the Sure-Thing Principle (Esponda & Vespa, 2016).

The main contributions of my paper are as follows.

(1) RSD outperforms TTC. The theory of obvious strategy-proofness can explain more dominant strategy play in OSP-RSD compared to TTC, but it cannot explain more dominant strategy play in SP-RSD compared to TTC.

(2) OSP-RSD does not outperform SP-RSD. Based on a new method of comparing simultaneous with sequential mechanisms, I find that there is no difference in dominant strategy play between the obviously strategy-proof and the strategy-proof version of RSD.

(3) The capability for contingent reasoning predicts dominant strategy play. Based on two additional games, I find that subjects with a perfect ability for contingent reasoning, as well as subjects who rather refrain from contingent reasoning, play dominant strategies.

(4) Subjects have preferences over mechanisms. A fraction of 40% of the subjects strictly prefers one mechanism to the other.

The remainder of the paper is organized as follows. Section 2 describes the assignment mechanisms, their theoretical properties, and the resulting predictions for the experiment. The experimental design is described in Section 3. Section 4 provides the main results. Section 5 concludes.

2. THEORETICAL FRAMEWORK AND MECHANISMS

Consider the *house allocation problem* as a triple $\langle I, O, \succ \rangle$, where I is a set of agents, O is a set of objects, and $\succ$ are strict preference profiles. Let $|I| = |O|$.

An assignment is $\mu : I \to O$. I rely on the standard formulation of Hylland & Zeckhauser (1979).

Parametrization. In the experiment, four indivisible objects $O = \{a, b, c, d\}$ are assigned to four agents $I = \{1, 2, 3, 4\}$. Each agent is assigned exactly one object. The agents' strict preference profiles $\succ_i$ are common knowledge:

$\succ_1$	$\succ_2$	$\succ_3$	$\succ_4$
a	a	a	a
b	b	d	d
c	c	c	c
d	d	b	b

The preference profiles are aligned. Agents form pairs with identical preferences. The designed alignment resembles correlated preferences in real life. An exchange opportunity is salient for objects b and d and it is symmetric between agents.

Based on this environment, three different stochastic assignment mechanisms are compared in the experiment: the simultaneous version of RSD (SP-RSD), the sequential version of RSD (OSP-RSD), and top trading cycles with random initial endowments (TTC).

In this paper, I use standard theory and the behavioral theory of obvious strategy-proofness to test the institutional design of stochastic assignment mechanisms. The mechanisms are evaluated based on three criteria: Strategy-proofness, obvious strategy-proofness, and Pareto efficiency.

Strategy-proofness (SP) requires truthful preference revelation to be the weakly dominant strategy for every agent. In addition, the behavioral theory of obvious strategy-proofness (OSP) takes cognitive limitations into account: Recognizing an obviously dominant strategy does not require hypothetical thinking about the other agents' actions (Li, 2017). An assignment is Pareto-efficient if it is in the core, i.e., no coalition improvement is possible. Efficiency is defined for each group as the sum of earnings divided by the sum of earnings which would have been obtained under the Pareto-efficient assignment.

2.1. Simultaneous random serial dictatorship (SP-RSD)

In the simultaneous version of RSD (SP-RSD), all agents submit their entire lists of preferences. It works as follows (description from Y. Chen & Sönmez, 2002).

- Agents submit a full list of preferences over objects.
- Nature draws an order of the agents from a uniform distribution.
- The first agent is assigned her top choice.
- The second agent is assigned her top choice among the remaining objects.

  ⋮

- The last agent is assigned the remaining object.

By regarding agents' preferences, SP-RSD dominates the random assignment of objects (Erdil, 2014). SP-RSD is strategy-proof and Pareto-efficient (Zhou, 1990; Abdulkadiroğlu & Sönmez, 1998). SP-RSD is not obviously strategy-proof, because recognizing the weakly dominant strategy involves identifying the other agents' potential actions by hypothetical thinking (Li, 2017).

2.2. Top trading cycles with random endowments (TTC)

Now I introduce initial endowments to the house allocation problem. Initial random endowments for Gale's top trading cycles algorithm give rise to a housing market. The *housing market* (Shapley & Scarf, 1974) is a quadruple $\langle I, O, \succ, \eta \rangle$, where η is the initial endowment assignment added to the house allocation problem. TTC with strict preferences selects the unique core allocation of the housing market and coincides with the competitive equilibrium (Roth & Postlewaite, 1977). Therefore, this mechanism is also known as "core from random endowments" (Abdulkadiroğlu & Sönmez, 1998; Pathak, 2008). TTC with random endowments works as follows (description adapted from Abdulkadiroğlu & Sönmez, 1998).

- Agents submit a full list of preferences over objects.
- Nature draws an initial assignment from a uniform distribution.
- Step 1: Every agent points to the agent owning her most preferred object. Every object points to its owner. There is at least one cycle. A cycle is an order of agents $\{1,2\}$ where agent 1 points to agent 2 and agent 2 points to agent 1. Execute trades and remove cycles. Go to the next step if there are remaining agents.

⋮

- Step k: Every remaining agent points to the agent owning her most preferred object among the remaining objects. Objects point to their owners. Execute trades and remove cycles. Repeat until there are no remaining agents.

TTC obtaining the unique core allocation is strategy-proof (Roth, 1982) as well as individually rational, and Pareto-efficient (Ma, 1994). However, TTC is not obviously strategy-proof for markets with 3 or more agents because the opposing players' actions need to be taken into account to recognize the dominant strategy.

Proposition Li. (Li, 2017) *TTC with $n \geq 3$ is not OSP-implementable.*

TTC coincides with RSD if there are no existing endowments. TTC, SP-RSD, and OSP-RSD are strategy-proof and ex-post Pareto-efficient.

Theorem AS. (Abdulkadiroğlu & Sönmez, 1998) *RSD is the same lottery mechanism as TTC with random endowments.*

2.3. Pareto efficiency in the experiment

TTC, SP-RSD, and OSP-RSD are ex-post Pareto-efficient in the experimental environment under the standard assumption of dominant strategy play.

Proposition 1. *The core allocation under SP-RSD, OSP-RSD, and TTC is: One agent gets her first, two agents get their second, and one agent gets her third preference.*

Sketch of proof. According to Abdulkadiroğlu & Sönmez (1998), TTC and serial dictatorships are ex-post Pareto-efficient. Therefore, TTC, SP-RSD, and OSP-RSD are ex-post Pareto-efficient and obtain the same outcome distribution. Independent of the order of the random queue, the following outcome distribution is obtained for the agents $I = \{1,2,3,4\}$.

1	2	3	4
a	a	a	a
b	b	d	d
c	c	c	c
d	d	b	b

Boxes around the preference profiles illustrate that only one agent gets her top choice object a, two agents get their second-best objects b and d, one agent receives her third preference c, and no agent receives her last preference. The outcome distribution of the core from mechanisms TTC, SP-RSD, and OSP-RSD, given the induced preference profiles $\succ_i$, is $1-2-1-0$. □

Since the preference profiles are highly correlated in the experimental parametrization, there is little room for differences in efficiency resulting from non-dominant strategy play.

2.4. Predictions

Since TTC, SP-RSD, and OSP-RSD are strategy-proof in standard game theory, subjects are predicted to play dominant strategies under all three mechanisms. The fraction of dominant strategies is expressed by using subscript s.

Hypothesis 1. (Dominant strategy play) $TTC_s = SP\text{-}RSD_s = OSP\text{-}RSD_s$,

where s=100% for all mechanisms.

Since OSP-RSD is obviously strategy-proof, subjects are predicted by the behavioral theory of OSP to play the dominant strategy more often under OSP-RSD than under SP-RSD or TTC under the assumption of limited ability for contingent reasoning. The fraction of dominant strategy play (subscript s) is predicted by the following relation:

Hypothesis 2. (Behavioral) $SP\text{-}RSD_s < OSP\text{-}RSD_s \wedge TTC_s < OSP\text{-}RSD_s$.

Hypotheses 1 and 2 are competing hypotheses about dominant strategy play from standard and behavioral theory.

Since the mechanisms are ex-post Pareto-efficient, subjects are predicted to obtain the core allocation under all three mechanisms. Proposition 1 predicts that this core allocation assigns one agent her top choice, two agents their second-best choice, and one agent her third choice. The fraction of core allocations is expressed by using subscript p, where p=100% in standard theory.

Hypothesis 3. (Pareto efficiency) $TTC_p = SP\text{-}RSD_p = OSP\text{-}RSD_p$.

3. EXPERIMENTAL DESIGN

Figure 1 summarizes the experimental setup required when implementing a one-shot game under complete information. The three mechanisms, TTC, SP-RSD, and OSP-RSD, are compared between-subjects. In the beginning, participants are randomly divided into groups of 4, where each participant is randomly assigned a role within each group. Roles correspond to preference profiles $\succ_i$ and remain constant. Then, subjects are informed about two mechanisms, one after another. Subjects only get to know mechanism 2 after they have completed mechanism 1. They are either confronted with the pair TTC/ SP-RSD or with the pair TTC/ OSP-RSD. Subjects act through each allocation rule and answer control questions. Instructions are available in the Appendix.

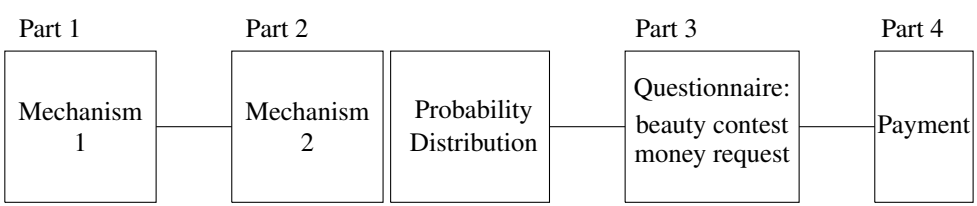

Figure 1: Experimental setup.

Part 1. Subjects state their preferences under mechanism 1. This one-shot decision is the main variable of interest for the between-group comparison. Mechanisms are either TTC, SP-RSD, or OSP-RSD. The mechanism is determined randomly. Subjects do not receive feedback until part 4. TTC and SP-RSD require submission of a list of preferences over 4 objects. In OSP-RSD, subjects select an object from a set of objects.

OSP-RSD is implemented using the strategy method (Selten, 1967): Subjects choose their preferred object 4 times from 4 sets of objects at the same time. The aim is to make the simultaneous submission of a list in TTC and SP-RSD comparable to selecting objects from sets in OSP-RSD. The reason is that listing 4 objects in an order may produce larger error rates than selecting a single object. Therefore, each subject is confronted with 4 sets of objects when making their choice in OSP-RSD. Subjects are ordered sequentially, but they do not know in which position of the sequence they are when they decide about the 4 sets.

Figure 2 illustrates the implementation of OSP-RSD with an example. Subjects face 4 choices in 4 different sets containing 4, 3, 2, and 1 object(s), without knowing their position in the sequence. Only one set of objects is the payoff-relevant set depending on the others' previous choice (corresponding to the actual position in the sequence). For instance, the subject at position 1 chooses her preferred object from sets $\{a,b,c,d\}$, $\{b,c,d\}$, $\{b,c\}$, and $\{b\}$, without knowing that she is at position 1. Her payoff-relevant decision is about the object from $\{a,b,c,d\}$.

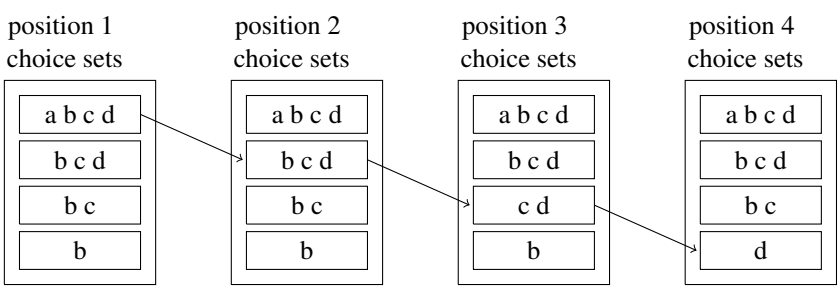

Figure 2: OSP-RSD: Every subject chooses from 4 choice sets without knowing her actual position in the sequence. The decision of the subject at position 1 in the first choice set impacts the second choice set of agent at position 2, etc.

This method provides a fair comparison of dominant strategy play between submitting a preference list and directly choosing from sets of objects because the task of listing 4 objects can be regarded as choosing an object for each individual rank.

Importantly, the 4 choices in OSP-RSD do not involve contingent reasoning about others' actions if we assume narrow bracketing (Rabin & Weizsäcker, 2009). Under narrow bracketing, assuming that subjects evaluate each choice

separately, the elicitation procedure remains OSP. That is, I presume that subjects choose their preferred object from set 1, then they choose their preferred object from set 2 and so on without considering the complete strategy plan. In this way, each element in the strategy plan is OSP.

The single choice according to the actual position in the random sequence, as used by Li (2017), can be reconstructed using this method by only considering the decision about choice set 1 by position 1, the decision about choice set 2 by position 2, and so on. Henceforth, I refer to this reconstruction as OSP-RSD(1).

Part 2. In order to elicit a preference over mechanisms, subjects also make a choice under mechanism 2. Mechanism 2 is either SP-RSD or OSP-RSD if mechanism 1 was TTC, or it is TTC if mechanism 1 was SP-RSD or OSP-RSD. This order is randomized to control for order effects.

Only one mechanism (either 1 or 2) is paid out randomly in the end to avoid hedging possibilities. I elicit the preference for a mechanism by giving subjects the possibility of determining the probability distribution for the random payment. They do not know about the distributions until they finish mechanism 2. They choose between 50:50 and 80:20, respectively, for each mechanism. The latter options, implying 80% probability of the preferred mechanism to be paid out, cost ten euro cents. An indifferent subject would choose 50:50, whereas a subject with a strict preference would choose the option 80:20, respectively.

Part 3. The questionnaire contains two games to capture extreme forms of contingent reasoning. In the *2-person beauty contest game* (Grosskopf & Nagel, 2008), two players form a group and submit an integer guess between 0 and 100. The closest guess to $2/3$ of the average of the group wins 2 euros. The weakly dominant strategy is stating 0. I presuppose that stating 0 is a sufficient condition for perfect contingent reasoning about the hypothetical guess of the other person.

In contrast, the *11-20 money request game* of Arad & Rubinstein (2012) does not have a pure strategy Nash equilibrium, but an intuitively salient level-0 strategy of requesting 20. In this game, two players form a group and submit an integer number between 11 and 20 points. They keep this amount. If one player submits exactly one integer less than the other player, then this player receives 20 points on top. The benefit of this second game is that answering 20 does not necessarily involve thinking about the others' hypothetical actions, and it is at the same time a sensible answer. Submitting 19 would imply that a player

thinks that the other player submits 20, and so on. In turn, submitting 20 would imply that she does not take into account the hypothetical action of the other player. Otherwise she would see that she could be better off by submitting one point less. However, I assume that any form of contingent reasoning would involve thinking about the other players' action expressed by submitting points < 20. Having a preference for maximum welfare could potentially also play a role when submitting 20, since the maximum number of points is achieved if one player states 20 and the other one states 19. Additionally, the psychological need for a cognition scale (Cacioppo & Petty, 1982; Bless et al., 1994), risk attitude measurement (Holt & Laury, 2002), and demographic characteristics are included.

Part 4. After the experiment, the uncertainty is resolved by running the mechanisms. Subjects are paid out one mechanism with the probability determined in part 2. They receive 10 euros for their top choice, 7 euros for their second, 4 euros for their third, and 1 euro for their least preferred object. Payments are administered anonymously and privately.

Procedure. Sessions were conducted in March 2017 at the laboratory of the Technical University of Berlin. In sum, $N = 228$ participants took part. They were on average 26 years old and 52% of them were female. The majority of subjects (57%) were students of engineering, mathematics, or physics. Sessions lasted around 60 minutes. Average earnings (including show-up fee) were 16 euros (min: 9; max: 22). The experiment was programmed using the software z-Tree (Fischbacher, 2007). Participants were recruited using ORSEE (Greiner, 2015).

4. RESULTS

4.1. Dominant strategy play

Do individuals play their dominant strategies? This question is analyzed based on the subjects' stated preferences in mechanism 1 (where they do not know about the second mechanism). Hypothesis 1 states that subjects play dominant strategies under all mechanisms. This paper does not find supporting evidence. The behavioral Hypothesis 2 states that subjects play dominant strategies more often in OSP-RSD than in TTC. This is supported by the data.

Result 1. (Dominant strategies: RSD versus TTC) *On average, dominant*

strategies are played more often in RSD mechanisms than in TTC. Being in the TTC mechanism decreases the likelihood of playing the dominant strategy by 30% compared to SP-RSD.

Support. Table 1 presents proportions of dominant strategy play under each mechanism. Subjects play dominant strategies significantly more often in the RSD mechanisms than in TTC. Table 4 presents probit estimation results of dominant strategy play. The marginal effect of TTC compared to SP-RSD (row 2) is significant and robust to controlling for additional explanatory factors in model (4).

Table 1: Proportion of dominant strategy play.

Mechanism	Obs.	Dominant strategies	Mann-Whitney U test p-value	
			vs. TTC	vs. SP-RSD
TTC	23	53.3%		
SP-RSD	18	86.1%	0.0002	
OSP-RSD(1)	16	95.3%	< 0.0001	0.029
OSP-RSD(4)	16	78.1%	0.0021	0.813

The average of each group is one independent observation. Reported p-values are based on one-tailed testing, except for SP-RSD vs. TTC. For TTC and SP-RSD, dominant strategy play is defined as submitting the full preference list truthfully. For OSP-RSD(1), dominant strategy play is defined as choosing the largest prize according to the position in the queue. For OSP-RSD(4), it is defined as choosing each time the largest prize out of 4 different sets with 4, 3, 2, and 1 alternative(s).

The behavioral Hypothesis 2 also states that subjects play dominant strategies more often in OSP-RSD than in SP-RSD. Employing two different definitions of OSP-RSD, labeled OSP-RSD(1) and OSP-RSD(4), this paper finds mixed evidence.

Both OSP-RSD definitions are obtained from the same choice data and method. OSP-RSD(1) is nested in OSP-RSD(4). In OSP-RSD(1) – consistent with the definition in Li (2017) – participants play a dominant strategy if they choose the largest prize according to their actual position in the random queue. In OSP-RSD(4), dominant strategy play is defined as choosing the largest prize in each and every set of objects simultaneously. Here, subjects make 4 choices

from 4 different sets containing 4, 3, 2, and 1 object(s).

Result 2. (Dominant strategies: OSP-RSD versus SP-RSD) *On average, subjects play their dominant strategies more often in OSP-RSD(1)* (95%) *than in SP-RSD* (86%). *There is no significant difference between OSP-RSD(4) and SP-RSD.*

Support. Table 1 presents proportions of dominant strategy play in each mechanism. Dominant strategy play in OSP-RSD(1) occurs significantly more often than in SP-RSD. Dominant strategies are not played significantly more frequently in OSP-RSD(4) than in SP-RSD.

Result 2 indicates that the difference between SP-RSD and OSP-RSD depends on the definition of dominant strategy play in OSP-RSD. The OSP-RSD(1) finding replicates previous results from Li (2017). Here, dominant strategy play is based on one single choice from one set of objects. However, the difference between OSP-RSD and SP-RSD vanishes when 4 choices are considered in OSP-RSD(4).

Why then is the frequency of dominant strategy-play lower in OSP-RSD(4) compared to OSP-RSD(1)? The reason is that the misrepresentation or error rate is amplified across multiple choice sets. OSP-RSD(4) defines a subject as misrepresenting if any of her four choices are not dominant strategies. In OSP-RSD(4), 5/14 cases of non-dominant strategy play can be attributed to selecting the second-best object from the full choice set containing 4 objects and at the same time playing the dominant strategy in the choice set containing 3 objects. Another 7/14 cases can be attributed to selecting the second-best object in the choice set containing 3 objects, while playing the dominant strategy in the full choice set with 4 objects.

4.2. Misrepresentation strategies

Which misrepresentation strategies are mainly used? Figure 3 presents an overview of the manipulation strategies in SP-RSD and TTC.[3] Only preference lists from mechanism 1 are considered only (one-shot, between-subjects).

[3] In OSP-RSD(1), all three cases of manipulation are second-best on top. In the experiment, four subjects use more than one of the three manipulation strategies. Then, the first strategy is taken into account according to the order: second on top – last on top – third on second.

The most common manipulation strategy is to put the 2nd preference at the top of the list. This *second-best on top* strategy accounts for 26/43 cases of non-dominant strategy play in TTC, for 3/10 in SP-RSD, for 3/3 in OSP-RSD(1), and for 12/14 cases of non-dominant strategy play in OSP-RSD(4).

Figure 3 illustrates the tendency that the *second-best on top* strategy is played more frequently in TTC than in SP-RSD. This indicates that it is easier to recognize for subjects to recognize that they cannot outsmart the mechanism by putting their second-best on top of the list in SP-RSD.

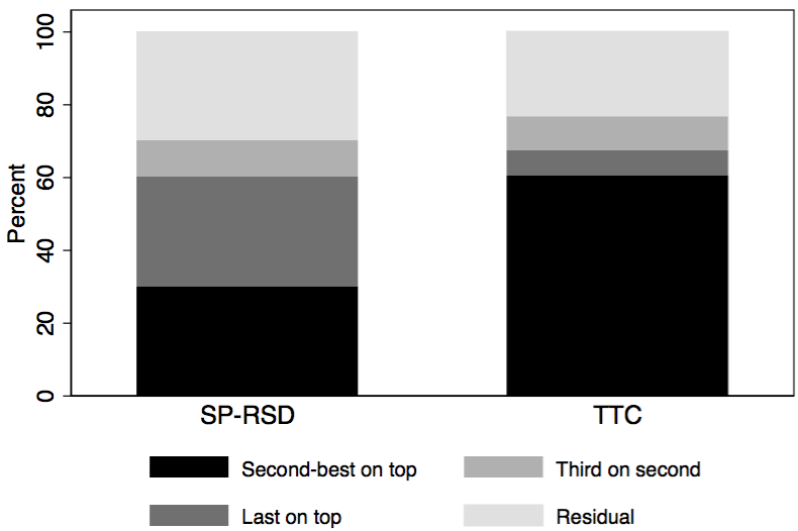

Figure 3: Misrepresentation strategies in SP-RSD and TTC.

The other two strategies of putting the *last preference on top* of the list and of putting the *third preference on the second* position are less frequent.

4.3. Efficiency

Does not playing the dominant strategy affect the efficiency of the outcome? Efficiency is defined for each group as the sum of payoffs divided by the sum of Pareto-efficient earnings. Observed efficiency is based on one random queue. Expected efficiency is based on simulations with 10.000 random queues. Hypothesis 3 states that all mechanisms result in the Pareto-efficient assignment. This is not supported by the evidence.

Result 3. (Pareto-efficient assignments) *In expectation, SP-RSD attains the Pareto-efficient assignment* 30% *more often than TTC.*

Support. Table 2 presents the proportions of assignments ending in the Pareto-efficient welfare level using the simulated expected efficiency in the final column. The expected proportion of Pareto-efficient assignments is significantly different between SP-RSD (72.2%) and TTC (39.1%), according to the Mann-Whitney U test ($N = 41$, $p = 0.037$, two-tailed).

Table 2: Efficiency.

Mechanism	Observed efficiency mean	Expected efficiency	
		Mean from 10.000 random draws	Proportion of Pareto-efficient assignments
TTC	0.953	0.943 (0.016)	39.1%
SP-RSD	0.976	0.964 (0.018)	72.2%
OSP-RSD(1)	0.960	—	—
Random lists	—	0.821 (0.026)	—

The Pareto-efficient welfare level under dominant strategy play is 1. Standard errors in parenthesis. Observed efficiency is based on one random queue. Expected efficiency is based on simulations with 10.000 random queues.

OSP-RSD(1) cannot be evaluated based on expected efficiency. Based on observed efficiency, the proportion of Pareto-efficient assignments in OSP-RSD(1) is 81.3%. This indicates that more dominant strategy play leads to more groups attaining the Pareto-efficient welfare level. Due to the correlated experimental environment, the mean efficiency does not differ between mechanisms (see Section 2.3).

4.4. Mechanisms

Why do individuals not play dominant strategies? I explore the potential explanations including the subjects' ability for contingent reasoning, procedural preferences for the mechanisms, and risk aversion.

Contingent reasoning. The capability to reason contingently is captured with data from two additional games. I focus on identifying the extremes.

Perfect *contingent* reasoning is defined as playing the dominant strategy in the 2-person beauty contest game by choosing 0. *Non-contingent* reasoning is defined as playing the level-0 strategy in the 11-20 money request game by requesting 20 and thereby refraining from reasoning contingently. Sample distributions of guesses and requests can be found in the Appendix.

Result 4. (Contingent reasoning) *Individuals with a perfect capability to reason contingently, as well as individuals who would rather refrain from reasoning contingently, play dominant strategies. The marginal effect of refraining from contingent reasoning on dominant strategy play is* 14%.

Support. Table 3 presents cases of dominant strategy play by the subjects' contingent reasoning. Subjects ex-ante classified as having the perfect capability, as well as subjects classified as refraining from reasoning contingently, play dominant strategies. Table 4 presents probit estimation results of dominant strategy play. Non-contingent reasoning (row 7) is significantly positively related to playing the dominant strategy. The marginal effect of non-contingent reasoning, compared to those not requesting 20 in the 11-20 money request game, is obtained from model (3).

Table 3: Ability for contingent reasoning and dominant strategy play.

Ability	Definition	Dominant strategies	Total cases
Contingent reasoning	Beauty contest guess $= 0$	14	14
Non-contingent reasoning	Money request $= 20$	19	21
Intermediate level	Beauty contest guess $> 0 \wedge$ Money request < 20	138	192
Residual	Beauty contest guess $= 0 \wedge$ Money request $= 20$	1	1
Total		172	228

Result 4 indicates that both extreme expressions of the ability for contingent reasoning result in dominant strategy play. This means that intermediate levels of the ability for contingent reasoning matter for dominant strategy play.

Table 4: Probit regression of dominant strategy play.

			Predicted variable: Dominant strategy play			
			(1)	(2)	(3)	(4)
1	Mechanisms:	OSP-RSD(1)	0.142*	0.082	0.087	0.081
	(Reference: SP-RSD)		(0.080)	(0.091)	(0.086)	(0.092)
2		TTC	-0.293***	-0.304***	-0.304***	-0.295***
			(0.071)	(0.068)	(0.070)	(0.065)
3	Preference for:	SP-RSD		-0.267***	-0.246***	-0.253***
	(Reference: Indifferent)			(0.089)	(0.089)	(0.088)
4		OSP-RSD		0.118*	0.119*	0.123
				(0.070)	(0.071)	(0.081)
5		TTC		0.040	0.027	0.005
				(0.067)	(0.069)	(0.067)
6	Beauty contest guess				-0.002*	-0.002*
					(0.001)	(0.001)
7	Non-contingent				0.137**	0.240**
					(0.064)	(0.108)
8	Risk aversion				-0.001	0.007
					(0.014)	(0.015)
9	Non-contingent ×					-0.043
	Risk aversion					(0.068)
10	Need for cognition					0.001
						(0.002)
11	Set of controls +		No	No	No	Yes
	Clusters		57	57	57	57
	N		228	228	228	228

Predicted variable: 1 if dominant strategy play and 0 otherwise. Data from mechanism 1.
Average marginal effects reported. Standard errors clustered at group level in parentheses.
Non-contingent: 1 if request is 20 in the 11-20 money request game and 0 otherwise.
+ Set of controls: demographics, player role, field of study, math grade.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4 also shows that the level of contingent reasoning, approximated by the beauty contest guess (row 6), is related to dominant strategy play. Note that the definition of perfect contingent reasoning cannot be included in the probit model because it perfectly predicts dominant strategy play.[4] The self-reported psychological need for cognition (row 10) is not significantly related to dominant strategy play.

Procedural preferences. Do subjects have a preference for mechanisms? Overall, around 40% of the subjects strictly prefer one assignment mechanism to the other, while the majority of participants is indifferent between the mechanisms. The TTC mechanism is preferred by 18%, SP-RSD by another 18%, and OSP-RSD by 22% of the subjects. This indicates that a substantial fraction of subjects has a preference for an assignment mechanism. Moreover, the preference for SP-RSD is found to be related to dominant strategy play.

Result 5. (Procedural preferences) *The likelihood of playing a dominant strategy is 25% lower for individuals having a preference for SP-RSD compared to indifference.*

Support. Table 4 presents probit estimation results of dominant strategy play. Having a preference for SP-RSD (row 3) is associated with a significantly decreased likelihood of dominant strategy play compared to the reference group of being indifferent. The marginal effect from the full model (4) is reported.

Result 5 suggests that having a procedural preference for SP-RSD is negatively related to dominant strategy play. More specifically, this effect can be partly attributed to subjects preferring the SP-RSD mechanism and not playing the dominant strategy in TTC. In 12/15 of the overall cases of having a preference for SP-RSD and not playing the dominant strategy, subjects do not play the dominant strategy in TTC.

Risk attitudes. Table 4 indicates that there is no significant relation at the 5% significance level between risk attitudes (row 8) and dominant strategy play.

[4] Contingent reasoning is defined as playing the weakly dominant strategy in the 2-person beauty contest game by choosing 0. However, choosing 0 in the beauty contest game perfectly predicts dominant strategy play in the mechanisms. This yields complete separability and the failure to fit the estimation models in Table 4 by maximum likelihood.

Risk aversion interacted with non-contingent (row 9) is not significantly related to dominant strategy play either. This indicates that the relation between not engaging in hypothetical thinking as measured by non-contingent reasoning (row 7) and dominant strategy play cannot merely be explained by subjects being risk-averse.

5. DISCUSSION AND CONCLUSION

In this paper, I test the performance of RSD and TTC for the house allocation problem *without* endowments and find that – in contrast to standard theory – RSD outperforms TTC. Dominant strategies are played more frequently and the Pareto-efficient welfare level is attained more often in RSD. This result stands in contrast to previous findings of Y. Chen & Sönmez (2002, 2004) on the house allocation problem with existing endowments. Their opposite finding that TTC outperforms RSD mainly results from subjects choosing the option not to participate in RSD because they could be worse off than their endowment matching. In line with the theory of Abdulkadiroğlu & Sönmez (1998), this option is not available in the present problem without endowments since everyone participates.

In the experiment, I introduce a new method to compare simultaneous with sequential mechanisms. I assume that listing objects in a preference order involves a decision about each individual rank in the list. Therefore, submitting a list (in TTC and SP-RSD) is more comparable to choosing from multiple choice sets than from a single choice set as in Li (2017). Sequential OSP-RSD is implemented as subjects choosing their preferred object from 4 choice sets without knowing their position in the sequence. This method has a second virtue: The single choice according to the actual position in the sequence can be reconstructed from this data. The single choice from one set is nested in the 4 choices from 4 sets. The new method comes at a cost: The decision situation of choosing simultaneously from different sets is cognitively more complex than choosing from one set. Nevertheless, reconstructing the single choice, I replicate the finding of Li (2017) that 2.6% (my experiment: 4.7%) do not play dominant strategies in OSP-RSD.

The methodological drawback of this new method is that the sequential version of RSD is only obviously strategy-proof if the narrow bracketing assumption (Rabin & Weizsäcker, 2009) is fulfilled: I assume that subjects evaluate each choice set separately when playing the strategy method. If narrow

bracketing is not fulfilled, identification could potentially be confounded since considering all four choices could still require contingent reasoning. Thus, when using the strategy method for obviously strategy-proof mechanisms, one has to weigh this drawback against the greater comparability between simultaneous and sequential mechanisms.

I do not find a difference in dominant strategy play between the sequential OSP-RSD and the simultaneous SP-RSD when using the data from the new method. The reason is that the error rate is higher when choosing from multiple sets: Subjects play the dominant strategy by choosing the highest prize in the first choice set, but fail to do so in the second choice set (and the other way around).

A substantial fraction of subjects (40%) prefers one assignment mechanism over the other, while the majority is indifferent. This evidence is in line with the previous finding that preferences over mechanisms which yield identical expected outcomes can differ systematically (Schmelzer, 2016). Further, the preference for SP-RSD is negatively related to dominant strategy play. This is driven by individuals who do not play the dominant strategy in TTC, indicating a preference for RSD.

I provide a first approach of identifying extreme forms of contingent reasoning. Individuals classified as having a perfect ability, as well as subjects who would rather refrain from contingent reasoning, play dominant strategies. This implies that the behavior of individuals with an intermediate level of contingent reasoning is crucial for the performance of mechanisms. For intermediate levels, mechanisms become key if the dominant strategy is easier to see than in others. In a related paper, Basteck & Mantovani (2018) find that subjects with a low cognitive ability play the dominant strategy less frequently in the deferred acceptance algorithm than high-ability subjects. The relation between the individuals' capability for contingent reasoning and their cognitive ability in matching markets remains an open question.

Several explanations for preference misrepresentation are discussed in the matching literature (Hassidim et al., 2017). In my experiment, the most prevalent form of misrepresentation is to place the second-best object at the top of the list; even more frequently in TTC than in SP-RSD. The top-choice object is identical for all subjects in a group while the second-best is only identical for half of the group. This misrepresentation behavior is consistent with the self-selection explanation of L. Chen & Pereyra (2016) that subjects rank an object lower if the perceived chance of receiving it is very low.

The findings of my paper can inform behavioral theory on obvious strategy-proofness (OSP). The strategic complexity of TTC and RSD seems to play a role, but in a different way than predicted by the behavioral theory of OSP. While OSP predicts the result that OSP-RSD outperforms TTC, it cannot explain the result that SP-RSD outperforms TTC. SP-RSD and TTC may differ in their complexity in a way so far not captured by OSP. A refinement of contingent reasoning might be helpful. The number or the quality of the contingencies may play a role in explaining why the frequency of dominant strategy play is larger in SP-RSD than in TTC. TTC not only involves contingent reasoning about the random priorities, as SP-RSD does, but also includes the mutually beneficial exchange opportunities given the initial random assignment.

From a policy perspective, the social planners' choice matters since the strategy-proofness of one mechanism is easier to understand than the strategy-proofness of the other. Given the experimental evidence, it is easier for individuals to recognize that they cannot game the system by misrepresenting their preferences in RSD than in TTC. As a consequence, RSD outperforms TTC. Therefore, the RSD mechanism may be considered instead of the TTC mechanism for assignment problems in which individuals' preferences are correlated and in which the individuals do not have pre-existing claims upon the objects to be assigned.

In conclusion, TTC and RSD are not equivalent as predicted by standard theory. In the absence of existing endowments, RSD yields more dominant strategy play than TTC. Dominant strategy play is related to the ability of individuals for contingent reasoning. Individuals with extremely high and low levels of contingent reasoning play dominant strategies. This can inform market design in practice: Strategy-proof and optimal assignment mechanisms may not yield the predicted results if individuals are boundedly rational. What is more, one strategy-proof market design may perform better with real people than the other.

Appendix

Figures

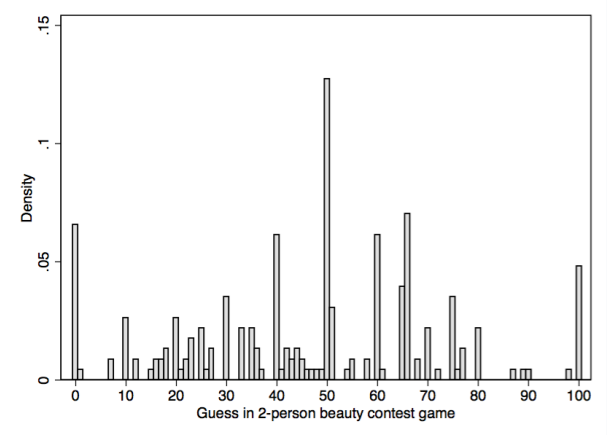

Figure 4: Distribution of guesses in the 2-person beauty contest game.

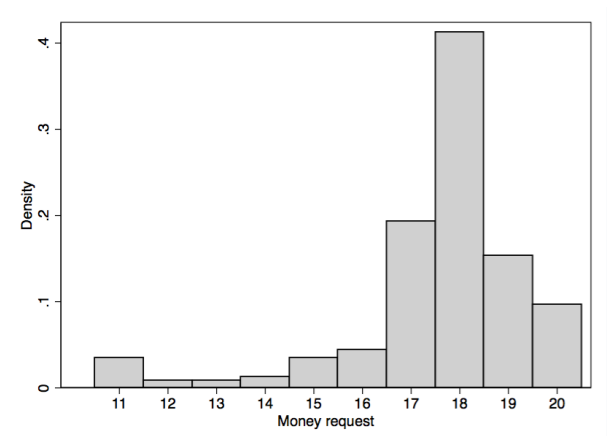

Figure 5: Distribution of requests in the 11-20 money request game.

Instructions

Welcome! You are about to take part in an economic study in decision-making. You will receive a show-up fee of 6 euros. Additionally, you will be able to earn a substantial amount of money. It is therefore crucial that you read these explanations carefully. The present instructions are identical for all participants.

Please switch off your mobile phone and do not communicate with other participants. If you have any questions, please raise your hand. We will then come over to you. Any violation of these rules means you will be excluded from the experiment and from any payments.

During the experiment, we will calculate in points. The total number of points you earn in the course of the experiment will be transferred into euro at the end, at a rate of

1 euro = 20 points.

The procedure and payment details are described below.

At the beginning of the experiment, all participants are randomly divided into groups of four. You will not get to know the identity of the other participants in your group. You stay in the same group during the experiment.

In the experiment, we simulate procedures that assign positions to applicants. A central clearinghouse takes care of the assignment procedure. You and the other participants are applicants. Within each group, you are randomly assigned the role of an applicant. This role remains the same throughout the experiment.

The following payment table determines your payoff at the end of the experiment.

Payment table.

Points	Applicant green	Applicant blue	Applicant red	Applicant yellow
200 points	W	W	W	W
140 points	X	X	Z	Z
80 points	Y	Y	Y	Y
20 points	Z	Z	X	X

In this payment table, you can see how many points each applicant receives for each assigned position. This table is equivalent in both tasks. For instance, if applicant green is assigned

position W in the allocation procedure, then he receives 200 points. If applicant green is assigned position X, then he receives 140 points; for position Y he receives 80 points and for position Z he receives 20 points.

Procedure

- Task 1
- Task 2
- Questionnaire
- Payment

You will receive more detailed information about task 1 and task 2 on your computer screen after the experiment starts. In each task, we simulate a procedure that assigns positions to applicants.

One of both tasks (either 1 or 2) is randomly determined and paid out to you in the end. The precise probability for the random payment will be determined in the experiment.

only for TTC and SP-RSD:

<< **Your Decision**

In each task, you will make a decision about a ranking of positions. You will receive details on the use of this ranking in the procedure during the experiment. All four applicants submit a ranking. You may submit any ranking. All positions have to be listed. Rank 1 means the top rank, rank two the second-highest, rank 3 the third-highest, and rank 4 the lowest rank. >>

Do you have any questions? If this is the case, then please raise your hand. We will answer your questions individually. Thank you for participating in this experiment!

Instructions for SP-RSD

All applicants submit one **ranking** of positions. Then the procedure works as follows.

- **List of applicants**: A fair lottery determines a list of applicants. This means each applicant has an equal chance of becoming first, second, third, or fourth on this list.
- The first applicant on the list of applicants receives the position at rank 1 of his submitted ranking.

- The second applicant on the list of applicants receives the position with the highest rank among the remaining positions of his submitted ranking.
- The third applicant on the list of applicants receives the position with the highest rank among the remaining positions of his submitted ranking.
- The fourth applicant on the list of applicants is assigned to the remaining position.

An example

Consider for illustration purposes the following example. There are three applicants (gray, black, and white) and three positions (A, B, and C).

- **Rankings.** Assume that the applicants submitted the following rankings:

Applicant gray:	rank 1 = B	rank 2 = C	rank 3 = A.
Applicant black:	rank 1 = C	rank 2 = A	rank 3 = B.
Applicant white:	rank 1 = B	rank 2 = C	rank 3 = A.

Important: These sample rankings are chosen arbitrarily and only serve illustrational purposes. They provide no guidance for your decision-making in the experiment!

- **List of applicants.** Assume the following list of applicants:

black – gray – white

Please complete the following sentences.

- The first applicant on the list of applicants receives the position at rank 1 of his submitted ranking. That is, applicant black receives position ___ .
- The second applicant on the list of applicants receives the position with the highest rank among the remaining positions (A and B) of his submitted ranking. That is, applicant gray receives position ___ .
- The third applicant on the list of applicants is assigned to the remaining position. That is, applicant white receives position ___ .

Instructions for OSP-RSD

The procedure works as follows.

- **List of applicants**: A fair lottery determines a list of applicants. This means each applicant has an equal chance of becoming first, second, third, or fourth on this list.
- **Set of positions**: Available positions which have not yet been chosen.
- The first applicant on the list of applicants chooses one position from the full set of positions.
- The second applicant on the list of applicants chooses one position from the remaining positions in the set of positions.
- The third applicant on the list of applicants chooses one position from the remaining positions in the set of positions.
- The fourth applicant on the list of applicants receives to the remaining position.

An example

Consider for illustration purposes the following example. There are three applicants (gray, black, and white) and three positions (A, B, and C).

- **List of applicants.** Assume the following list of applicants:

  black – gray – white

- Assume, for instance, that applicant black prefers position A and applicant gray prefers position B.

Please complete the following sentences.

- The first applicant on the list of applicants chooses one position from the full set of positions (A, B, C). That is, applicant black chooses position ___ .
- The second applicant on the list of applicants chooses one position from the remaining positions in the set of positions (B, C). That is, applicant gray chooses position ___ .
- The third applicant on the list of applicants receives to the remaining position. That is, applicant white receives ___ .

Your Decision

In this task, you will see 4 different sets of positions corresponding to a place on the list of applicants. One set of positions is relevant for your final payment.

Only at the end of the experiment will you learn which set of positions is relevant for your payment and with that, which place on the list of applicants you have.

Your decision is to choose your preferred position out of each of the 4 sets of positions.

Instructions for TTC All applicants submit one **ranking** of positions. Then the procedure works as follows.

- **Tentative assignment**: Each applicant is first tentatively assigned to one position based on a fair lottery. This means each applicant has an equal chance to be assigned to a particular position. This assignment is tentative.
- Next, the rankings are used to determine mutually beneficial exchanges between two or more participants.
- **Queue**: In order to perform mutually beneficial exchanges, a queue is determined by a fair lottery. The lottery determines each applicant's place in the queue. Each queue is equally likely. This means that each applicant has an equal probability of becoming first, second, ..., or last in the queue.
- The specific allocation process is explained below. It starts with the first applicant in the queue. The application of the first applicant in the queue is submitted to the position with rank 1 on his ranking.
  - If the application is submitted to his tentatively assigned position, then his tentative assignment is finalized, i.e., he receives the position. The applicant and his assignment are removed from subsequent allocations. The process continues with the next applicant in line.
  - If the application is submitted to another position, say position S, then the first applicant in the queue who is tentatively assigned position S is moved to the top of the queue directly in front of the requester.
- Whenever the queue is modified, the process continues similarly: An application is submitted to the highest ranked position for the applicant at the top of the queue.

- A mutually beneficial exchange is obtained when a cycle of applications are made in sequence, which benefits all affected applicants, e.g., A applies to B's tentatively assigned position, and B applies to A's tentatively assigned position. In this case, the exchange is completed and the applicants as well as their assignments are removed from subsequent allocations.
- The process continues until all applicants are assigned a position.

An example

Consider for illustration purposes the following example. There are three applicants (gray, black, and white) and three positions (A, B, and C).

- **Rankings.** Assume that the applicants submitted the following rankings:

Applicant gray:	rank 1 = B	rank 2 = C	rank 3 = A.
Applicant black:	rank 1 = C	rank 2 = A	rank 3 = B.
Applicant white:	rank 1 = B	rank 2 = C	rank 3 = A.

Important: These sample rankings are chosen arbitrarily and only serve illustrational purposes. They provide no guidance for your decision-making in the experiment!

- **Tentative assignment.** Assume the following tentative assignment of positions:

Applicant gray	Applicant black	Applicant white
C	B	A

- **Queue.** Assume the following queue of applicants:

gray – black – white

Please complete the following sentences.

- The application of the first applicant in the queue is submitted to the position with rank 1 on his ranking. That is, the application of applicant gray is submitted to position ___ .
- This application is not submitted to his tentatively assigned position C. The queue is modified: The first applicant in the queue who is tentatively assigned position B is moved to the top of the queue directly in front of the requester. At the top of the queue is now applicant ___ .

- The queue is modified:

  gray – black – white

  The new queue is:

  black – gray – white

- The application of the (new) first applicant in the queue is submitted to the position with rank 1 on his ranking. That is, the application of applicant black is submitted to position ___ .

- This application of applicant black is not submitted to his tentatively assigned position B. Therefore, the queue is modified: The first applicant in the queue who is tentatively assigned position C is moved to the top of the queue directly in front of the requester. At the top of the queue is now applicant ___ .

- The queue is modified:

  black – gray – white

  The new queue is:

  gray – black – white

- Now a cycle of applications is made in sequence and with that a mutually-beneficial exchange is obtained. The following two applicants exchange their tentatively assigned positions and are removed with their assignments from subsequent allocations:

  Applicant ___ and

  Applicant ___ .

- Illustration of the exchange of positions:

  B

  gray black

  C

- Applicant white is the remaining applicant in the assignment process. His application is submitted to the remaining position: ___ .
- Since applicant white has already been tentatively assigned to this position, the assignment is finalized and he gets the position. Applicant white and his position are removed from the allocation process.
- The final assignment is:

Applicant gray	Applicant black	Applicant white
B	*C*	*A*

References

Abdulkadiroğlu, A., & Sönmez, T. (1998). Random serial dictatorship and the core from random endowments in house allocation problems. Econometrica, 66(3), 689–701.

Abdulkadiroğlu, A., & Sönmez, T. (1999). House allocation with existing tenants. Journal of Economic Theory, 88(2), 233–260. doi: 10.1006/jeth.1999.2553

Arad, A., & Rubinstein, A. (2012, December). The 11-20 money request game: A level-k reasoning study. American Economic Review, 102(7), 3561–73. doi: 10.1257/aer.102.7.3561

Ashlagi, I., & Gonczarowski, Y. A. (2016). No stable matching mechanism is obviously strategy-proof. Working paper, Stanford University.

Basteck, C., & Mantovani, M. (2018). Cognitive ability and games of school choice. Games and Economic Behavior, 109, 156–183. doi: https://doi.org/10.1016/j.geb.2017.12.011

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for cognition: eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben. Zeitschrift für Sozialpsychologie, 25.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. Journal of Personality and Social Psychology, 42(1), 116.

Charness, G., & Levin, D. (2009). The origin of the winner's curse: A laboratory study. American Economic Journal: Microeconomics, 1(1), 207–236.

Chen, L., & Pereyra, J. S. (2016). Self-selection in school choice. Working paper.

Chen, Y., & Sönmez, T. (2002). Improving efficiency of on-campus housing: An experimental study. American Economic Review, 92(5), 1669–1686.

Chen, Y., & Sönmez, T. (2004). An experimental study of house allocation mechanisms. Economics Letters, 83(1), 137–140. doi: 10.1016/j.econlet.2003.10.012

Erdil, A. (2014). Strategy-proof stochastic assignment. Journal of Economic Theory, 151, 146–162.

Esponda, I., & Vespa, E. (2014). Hypothetical thinking and information extraction in the laboratory. American Economic Journal: Microeconomics, 6(4), 180–202.

Esponda, I., & Vespa, E. (2016). Contingent preferences and the sure-thing principle: Revisiting classic anomalies in the laboratory. Working paper.

Fischbacher, U. (2007, June). z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics, 10(2), 171–178.

Glazer, J., & Rubinstein, A. (1996). An extensive game as a guide for solving a normal game. Journal of Economic Theory, 70(1), 32–42.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. Journal of the Economic Science Association, 1(1), 114–125.

Grosskopf, B., & Nagel, R. (2008). The two-person beauty contest. Games and Economic Behavior, 62(1), 93–99.

Guillen, P., & Hakimov, R. (2016). Not quite the best response: Truth-telling, strategy-proof matching, and the manipulation of others. Experimental Economics, 20(3), 670-–686.

Hassidim, A., Marciano, D., Romm, A., Shorrer, R. I., et al. (2017). The mechanism is truthful, why aren't you? American Economic Review, 107(5), 220–24.

Hassidim, A., Romm, A., & Shorrer, R. I. (2016). "Strategic" behavior in a strategy-proof environment. In Proceedings of the 2016 ACM conference on economics and computation (pp. 763–764). New York, NY, USA: ACM. doi: 10.1145/2940716.2940751

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. American Economic Review, 92(5), 1644–1655.

Hylland, A., & Zeckhauser, R. (1979). The efficient allocation of individuals to positions. Journal of Political Economy, 87(2), 293–314.

Li, S. (2017, November). Obviously strategy-proof mechanisms. American Economic Review, 107(11), 3257–87. doi: 10.1257/aer.20160425

Ma, J. (1994). Strategy-proofness and the strict core in a market with indivisibilities. International Journal of Game Theory, 23(1), 75–83.

Moraga, J. F.-H., & Rapoport, H. (2014). Tradable immigration quotas. Journal of Public Economics, 115, 94–108. doi: 10.1016/j.jpubeco.2014.04.002

Ngangoue, K., & Weizsäcker, G. (2015). Learning from unrealized versus realized prices. Working paper, DIW Berlin.

Olson, M., & Porter, D. (1994). An experimental examination into the design of decentralized methods to solve the assignment problem with and without money. Economic Theory, 4(1), 11–40.

Pathak, P. A. (2008). Lotteries in student assignment: The equivalence of queueing

and a market-based approach. Working paper.

Pycia, M., & Troyan, P. (2018). Obvious dominance and random priority. Working paper. doi: 10.2139/ssrn.2853563

Rabin, M., & Weizsäcker, G. (2009). Narrow bracketing and dominated choices. American Economic Review, 99(4), 1508–43.

Rees-Jones, A. (2017). Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match. Games and Economic Behavior, 108, 317–330. doi: 10.1016/j.geb.2017.04.011

Roth, A. E. (1982). Incentive compatibility in a market with indivisible goods. Economics letters, 9(2), 127–132.

Roth, A. E., & Postlewaite, A. (1977). Weak versus strong domination in a market with indivisible goods. Journal of Mathematical Economics, 4(2), 131–137.

Schmelzer, A. (2016). Single versus multiple randomization in matching mechanisms. Working paper, Max Planck Institute for Research on Collective Goods.

Selten, R. (1967). Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. In H. Sauermann (Ed.), Beiträge zur Experimentellen Wirtschaftsforschung (pp. 136–168). Tübingen: J.C.B. Mohr (Paul Siebeck).

Shapley, L., & Scarf, H. (1974). On cores and indivisibility. Journal of Mathematical Economics, 1(1), 23-37. doi: 10.1016/0304-4068(74)90033-0

Troyan, P. (2016). Obviously strategyproof implementation of allocation mechanisms. Working paper.

Zhang, J. (2016). Level-k reasoning in school choice. Working paper.

Zhang, L., & Levin, D. (2017). Bounded rationality and robust mechanism design: An axiomatic approach. American Economic Review, 107(5), 235–239.

Zhou, L. (1990). On a conjecture by Gale about one-sided matching problems. Journal of Economic Theory, 52(1), 123–135.

Journal of Mechanism and Institution Design
ISSN: 2399-844X (Print), 2399-8458 (Online)
DOI:10.22574/jmid.2018.12.003

EFFICIENT RANDOM ASSIGNMENT WITH CARDINAL AND ORDINAL PREFERENCES

James Fisher

United States Automobile Association (USAA), USA
jamescdfisher@gmail.com

ABSTRACT

We develop a finite random assignment model where players know either their cardinal or their ordinal preferences and may make cardinal or ordinal reports to an assignment mechanism. Under truthful reporting, we find that all mechanisms that disregard the cardinal information in players' reports (e.g., Deferred Acceptance and Probabilistic Serial) are utilitarian inefficient, as are classic mechanisms that make use of cardinal information (e.g., Pseudo-markets). Motivated by these negative results, we introduce a "Simple Mechanism" that makes use of cardinal information "in the right way." We establish that this mechanism is utilitarian efficient, treats equals equally, and makes truth-telling almost Bayesian incentive compatible.

Keywords: Cardinal and ordinal preferences, random assignment, utilitarian efficiency.

JEL Classification Numbers: C7, D7, D8.

1. INTRODUCTION

Consider the problem of assigning two equally qualified doctors, Ann and Bob, to one of two medical research fellowships, at UC San Francisco and Mount Sinai Hospital in New York. Ann has prior knowledge of the colleagues at

I am grateful to Isa Hafalir, Tim Flannery, Antonio Miralles, Marek Pycia, John Wooders, Tiemen Woutersen, Yilei Zhang, and several anonymous referees for helpful comments, discussions, and insights. This paper was partially written while at University of California Los Angeles and the University of Arizona; I am grateful for their hospitality. All remaining errors are my own.

each hospital, the location (and available activities), and the salary, and has concluded that her value for UC San Francisco is 2, while her value for Mount Sinai is 1. Bob, who has been traveling with Doctors without Borders for the last few years, only knows that he strictly prefers UC San Francisco to Mount Sinai. The question is: given these preferences, how should we assign the fellowships to maximize social surplus/utilitarian welfare (without using money)?[1]

Our aim is to answer questions of this type. The key insight of our approach is that Bob, like every other person, has a preference intensity, i.e., he can like UC San Francisco a bit more than Mount Sinai or he can like it a lot more. Thus, Bob's ordinal preference is actually induced by a true, underlying cardinal preference that, we suppose, he does not know. (We discuss the reasons why a person may only know their ordinal preferences below.)

Since Bob has an underlying cardinal preference, we are able to estimate it from his ordinal preference and are able to use this estimate to compute a surplus maximizing assignment. For instance, if he has three possible values for each hospital of 1, $\frac{3}{2}$, and 2, which are equally likely, then we estimate that his values for UC San Francisco and Mount Sinai are $\frac{11}{6}$ and $\frac{7}{6}$ respectively.[2] It follows that it is surplus maximizing to send Ann to San Francisco and send Bob to New York – this assignment generates an (expected) social surplus of $2+\frac{7}{6}=\frac{19}{6}$.

In this paper, we study the elicitation and estimation of preferences, as well as their use in constructing surplus maximizing assignments, in a simple random assignment model with a finite number of players N and a finite number of objects K. Each player may get at most one object and an assignment is a map that tells us which players get which objects; a mixed assignment is a lottery over assignments.

The players are Bayesians with Von-Neumann Morgenstern utilities. To capture the idea that some know their cardinal preferences and others only

[1] This paper contributes to the "random assignment literature," which seeks to understand how goods should be allocated when monetary transfers are not possible. This is the case, for instance, with dorm rooms, fellowships, organs, school seats, and public programs (e.g., day-care and low-cost housing) where resell is impossible and other transfers to prevent participation in the assignment and/or alter reports are illegal. (In fact, California's Cartwright Act would prohibit Bob from paying Ann to not apply to UC San Francisco.)

[2] There are nine equally likely possible pairs of values for UC San Francisco and Mount Sinai: $(1,1)$, $(1,\frac{3}{2})$, …, $(2,\frac{3}{2})$, and $(2,2)$. Of these, only $(\frac{3}{2},1)$, $(2,1)$, and $(2,\frac{3}{2})$ are compatible with Bob's ordinal preference. Averaging these three gives $(\frac{11}{6},\frac{7}{6})$.

know their ordinal preferences, we suppose that nature draws a vector of "true values" for each player from a finite set of possible true value vectors. Nature then randomly determines which players observe their true values and which observe the rankings implied by their true values.

To digress, there are a number of reasons a person may only observe her ordinal preferences. For example, it may be that the objects are experience goods and, absent prior consumption, she may only gain a general ranking from the reviews and recommendations of others: a classic example is restaurants, where a Zagat's guide or Yelp reviews only provide a general ranking and the actual dining experiences are what provide the true values; additional examples include colleges, housing, and the like. Alternatively, the objects may be complex and hard to evaluate – e.g., home mortgages, health insurance contracts, employment opportunities, retirement plans, and the like. Hence, a person may avail herself of the reviews and recommendations of others to gain a general idea of how the objects rank or, if she has access to the needed information and is sufficiently skilled, she may be able to determine her cardinal values.[3] We abstract away from specifics by assuming some players are endowed with knowledge of their cardinal preferences, while others are not.

After players learn their true values or rankings, i.e., their types, they send reports to a mechanism. The mechanism takes in these reports, processes them, and then returns a mixed assignment. Subsequently, an outcome of this assignment is realized and players receive their objects. We focus on direct mechanisms for simplicity.

We are primarily concerned with interim (utilitarian) efficiency. An assignment is interim efficient if it maximizes (expected) social surplus once players know their types. A mechanism is interim efficient if it selects an interim efficient assignment (when players report truthfully).[4] Interim efficiency implies interim and ex-post Pareto efficiency, as well as interim and ex-post individual rationality.

Within this framework, we first examine whether classical mechanisms –

[3] Introspection suggests that, as a result of general reviews and rankings, a person may only get an "imprecise" estimate of her preference intensity – e.g., she may only whether one object is a bit more valuable than another or a lot more valuable – which she is unable to refine to a point estimate. In contrast, a Bayesian player always has a point estimate of her expected preference intensity. Hence, we view the Bayesian nature of players as an analytically tractable approximation of real-world estimating behavior.

[4] We study incentive compatibility separately from efficiency.

e.g., Boston, Deferred Acceptance, Random Serial Dictatorship, Probabilistic Serial, and Pseudo-markets – produce efficient assignments. Some of these mechanisms make use of the cardinal information contained in players' reports (e.g., Pseudo-markets), while others "disregard" it (e.g., Boston, Deferred Acceptance, Random Serial Dictatorship, and Probabilistic Serial).

Our first principal result is a negative one: under broad conditions – roughly, that two different, possible vectors of true values induce the same ordinal preference – any mechanism that disregards cardinal information (i.e., that only uses the ranking implied by players' reports) is interim inefficient (see Proposition 1). This result strengthens the classical intuition that such mechanisms "may be inefficient" and implies that they only achieve the second best.[5] The result is grounded in the fact that, when a mechanism disregards cardinal information, objects do not go to those who value them most. (We develop this intuition and those related to our other results in detail in the Illustration of Section 3.) Subsequently, we investigate whether mechanisms that make use of cardinal information are interim efficient and find that inefficiencies occur (see Examples 1 and 2).

Motivated by these negative findings, we introduce a "Simple Mechanism." It takes in players' ordinal and cardinal reports and then constructs estimates of their true values by treating their reports as though they were their types. It next computes σ, the set of pure assignments that maximize the sum of players' estimated payoffs. Subsequently, it generates a mixed assignment by implementing each element of σ with equal probability.

Our second principal result establishes that the Simple Mechanism is interim efficient and "symmetric" (see Proposition 2). The intuition for efficiency is that, under truthful reporting, the mechanism correctly estimates (expected) true values and thus maximizes social surplus. For instance, if we run it in our fellowship problem, then it generates the efficient assignment – sending Ann to San Francisco and Bob to New York – because it estimates that she values San Francisco more than he does. The Simple Mechanism's symmetry ensures that "equals are treated equally," i.e., if two players make the same report and have the same type, then they have the same payoff. Symmetry follows from

[5] The magnitude of the inefficiencies can be substantial. For instance, if we run Random Serial Dictatorship for our fellowship problem, then Ann and Bob would both have a $\frac{1}{2}$ chance of ending up at either fellowship; social surplus in this assignment is only $(2+1)\frac{1}{2}+(\frac{11}{6}+\frac{7}{6})\frac{1}{2}=\frac{18}{6}$. Thus, the inefficiency is $\frac{1}{6}$, which is approximately 25% of the Ann and Bob's average value for the fellowships.

the mechanism's uniform randomization.

Since efficiency implies individual rationality, it follows that players are open to using the Simple Mechanism. The real question, then, is whether they report truthfully so that the mechanism achieves its designed ends.

Our third principal result tackles this question: we establish that the Simple Mechanism makes truth-telling (i.e., truthful reporting of one's type) "approximately" incentive compatible. Specifically, we prove that for every $\varepsilon > 0$, no player i gains more than ε by lying strategically about her type to the Simple Mechanism instead of truthfully reporting it when (i) all other players report truthfully and (ii) and the number of players N is large (see Proposition 3). The core intuition is that, holding the set of objects fixed, as N grows, the probability i can get any object decreases because of "competition" with the other players. Thus, i's payoffs to lying and to telling the truth converge to her value of being unassigned from above, implying her gain from lying is less than ε for large N.

While there are situations where supply of objects is fixed, it is often natural to think that it increases with the number of players – e.g., colleges offer more courses in response to rising enrollments. Our fourth principal result addresses this point: we establish that there is a way to grow the set of objects with N via replication, while maintaining approximate incentive compatibility (see Proposition 4). The key insight is that, as N grows, replication must happen sufficiently slowly so that competitive pressures are preserved.

Since strategic manipulation is often costly – in terms of effort and time – our results imply that the Simple Mechanism makes truth-telling a Nash equilibrium. This, along with efficiency and symmetry, suggest that the Simple Mechanism may be a good mechanism to employ in situations where there is substantial demand for objects – as might be the case in the assignment of prime university housing or courses, in the awarding of grants or fellowships, in the allocation of slots in programs for free/subsidized day-care or housing, in the apportionment of prime employee parking spots or offices, or even in the distribution of kidneys and other organs.

The rest of this section surveys the related literature. Section 2 describes the model. Section 3 illustrates our results via an example and discusses the intuition behind them. Section 4 gives Proposition 1 and Examples 1 and 2. Section 5 describes the Simple Mechanism and gives Propositions 2 to 4. All proofs are provided in the Appendix and the Online Appendix collects supplemental results.

RELATED LITERATURE

We make three contributions to the modern random assignment literature. First, we allow players to know either their ordinal or cardinal preferences. Second, we strengthen a classic intuition by showing that all mechanisms that disregard cardinal information must be interim inefficient. Third, and most importantly, we show that efficiency can be recovered in an approximately incentive compatible fashion by giving a new, Simple Mechanism.

Regarding the first contribution, in the random assignment literature either (i) all players know their own cardinal preferences (e.g., Bu (2014), Budish (2011), Hafalir & Miralles (2014), He et al. (2018), and Hylland & Zeckhauser (1979)) or (ii) all players know their own ordinal preferences (e.g., Abdulkadiroglu & Sonmez (1998), Bogomolnaia & Moulin (2001), Kasajima (2013), Liu & Pycia (2016), Nicolo & Rodriguez-Alvarez (2013), and Sonmez & Unver (2010)). We view this polarity as artificial – since, in practice, some people know more about their preferences than others – and we show one way to dispense with it.

Regarding the second contribution, there is a classic intuition in the random assignment literature: mechanisms that disregard cardinal information *may* be (interim) inefficient (e.g., Abdulkadiroglu et al. (2008)). We strengthen this intuition by giving broad conditions under which these mechanisms *must be* inefficient. In this capacity, our work is closely related to the work of Pycia (2011). Pycia establishes that any ordinal, symmetric, regular, Pareto efficient, and asymptotically strategy-proof mechanism is inefficient.[6] Our relative contribution is to remove many of these qualifiers and cleanly identify the source of the inefficiency.

Our (interim) inefficiency result complement those of Che & Tercieux (2013) and Zhou (1990). Che and Tercieux show that the ex-ante per-capita utilitarian inefficiency of Pareto efficient mechanisms converges to zero as the number of players (and objects) goes to infinity. In light of this, our inefficiency result implies that the ex-ante per-capita utilitarian inefficiency must be strictly positive outside of the limit when a mechanism disregards

[6] An ordinal mechanism is one that only accepts ordinal reports, while a regular mechanism is one where, when the number of players is large, no small group of players can "substantially impact the allocations of other..." players (Liu & Pycia, 2016, 3). In addition, a mechanism is asymptotically strategy-proof when, for each $\varepsilon > 0$, players do not gain for than ε from lying when (i) all other players tell the truth, (ii) all players' types are known, and (iii) the number of players is large.

cardinal information. Zhou identifies a parallel source of inefficiency: the interaction of strategy-proofness and symmetry. Together, our results enrich the literature's understanding of when inefficiencies occur.

Regarding the third contribution, the Simple Mechanism is most closely related to the work of Nguyen et al. (2014). Nguyen et al. develop a random assignment model where players can acquire more than one object. They assume that players know their cardinal preferences and develop a mechanism based on a linear programming problem that maximizes social surplus subject to the constraint that the assignment be envy-free. They show that their mechanism is (approximately) constrained efficient and is asymptotically strategy-proof when the programming problem has a *unique* solution. Our relative contribution is to give the Simple Mechanism, which is always unconstrained efficient and always has economically compelling truth-telling properties.[7]

The Simple Mechanism is further related to the work of Budish (2011). Budish studies (in passing) a "utilitarian" mechanism that solicits players cardinal reports and assigns objects to maximize utilitarian surplus, using an environment where cardinal preferences are known.[8] Unlike the Simple mechanism, Budish's mechanism does not randomize. He concludes that this mechanism is "manipulable" in the sense that a player may always have an incentive to lie; this conclusion extends to the Simple Mechanism (and we discuss it in the Illustration of Section 3). Our relative contribution is to show that the gains to manipulation shrink as competition intensifies due to growth in the number of players.

2. THE MODEL

To rigorously study efficient assignment, we need to pin down the economic environment and introduce a few definitions concerning efficiency and incentive compatibility.

Economic Environment

There is a finite set of players $\mathscr{N} = \{1, \ldots, N\}$ and a finite set of objects $\mathscr{O} = \{1, \ldots, K\}$, where N and K are positive integers. We write i for the i-th player and o for the o-th object. Each player may be assigned to an object or

[7] Note, however, that the Simple Mechanism may not be envy-free – e.g., in the efficient assignment of our fellowship example, Bob wants Ann's position.

[8] The thrust of Budish (2011) is the introduction of an "approximate Pseudo-market" mechanism, which is Pareto efficient and "almost envy-free," but need not be interim efficient or symmetric.

may be unassigned, and each object may be assigned to at most one player. We adopt the convention that a player is unassigned if she is assigned to the "null" item η and write $\mathscr{O}' = \mathscr{O} \cup \{\eta\}$. A **pure assignment** ϕ is a function that specifies each player's assignment, i.e., $\phi : \mathscr{N} \to \mathscr{O}'$ such that, for each $i \in \mathscr{N}$ and each $o \in \mathscr{O}$, $\phi(i) = o \iff \phi^{-1}(o) = i$, where ϕ^{-1} is the inverse of ϕ.

We write Φ for the finite set of all pure assignments and we index the elements of this set from 1 to $|\Phi|$, where $|\cdot|$ gives the cardinality of a set. Let $\Delta\Phi = \{(x_1, \ldots, x_{|\Phi|}) \in \mathbb{R}_+^{|\Phi|} | \sum_{j=1}^{|\Phi|} x_j = 1\}$ be the set of probability distributions on Φ. A $\tilde{\phi} = (\tilde{\phi}_1, \ldots, \tilde{\phi}_{|\Phi|}) \in \Delta\Phi$ specifies a lottery over pure assignments where the l-th pure assignment is realized with probability $\tilde{\phi}_l$; the lottery is independent of all other randomness. We refer to $\tilde{\phi}$ as a **mixed assignment**. The **support** of $\tilde{\phi} = (\tilde{\phi}_1, \ldots, \tilde{\phi}_j, \ldots, \tilde{\phi}_{|\Phi|})$ is $\{\phi_j \in \Phi | \tilde{\phi}_j > 0\}$, i.e., the set of pure assignments on which it places non-zero probability.

We have in mind that some players know their cardinal values for objects and for being unassigned, while other players only know their ordinal preferences. To model this, we suppose that each player has a "true (monetary) value" for each object and for being unassigned. We only allow certain players to observe their true values. The other players observe binary relations that are "consistent" with their true values. Nature randomly decides which players observe their true values. For simplicity, the true values are independently and identically distributed and each player has the same chance of observing her true values.[9]

Specifically, for each player i, let $v_i = (v_{\eta i}, v_{1i}, v_{2i}, \ldots, v_{Ki})$ denote the vector of i's true values over η and the objects in $\mathscr{O}$, where $v_{\eta i}$ is i's value of η, v_{1i} is i's value of object 1, v_{2i} is i's value of object 2, and so on. These values vectors are drawn by nature from a finite set $\mathscr{V} \subset \mathbb{R}^{K+1}$ according to probability mass function $f_v(v_i)$, where $f_v : \mathscr{V} \to (0, 1]$ such that $\sum_{v \in \mathscr{V}} f_v(v) = 1$; the draw is independent of all other randomness. These values are monetary in nature and are nontransferable (due to institutional, legal, ethical, or other constraints).

Let $\mathscr{B}$ be the set of all complete, transitive, and reflexive binary relations on $\mathscr{O}'$. We say that a $\preceq \in \mathscr{B}$ is **consistent** with a $v = (v_\eta, v_1, v_2, \ldots, v_K)$ if $v_o \leq v_{o'} \iff o \preceq o'$ and $v_o \not\leq v_{o'} \iff o \not\preceq o'$ for all o and o' in $\mathscr{O}'$. That is,

[9] In the Online Appendix, we show that our results generalize to an environment where players have correlated true values, as well as arbitrary (and possibly correlated) knowledge of their true values.

$\preceq$ is consistent with v if it is the ordinal preference of a decision maker whose true values are v. We also say that v is consistent with $\preceq$ if $\preceq$ is consistent with v. Let $\mathscr{B}' = \{\preceq \in \mathscr{B} | \preceq$ is consistent with a $v \in \mathscr{V}\}$ be the subset of $\mathscr{B}$ whose binary relations are consistent with the values in $\mathscr{V}$; we work principally with $\mathscr{B}'$.

Lemma 1. The Consistency Relationship.
For each $v \in \mathscr{V}$, there is a unique consistent $\preceq \in \mathscr{B}'(\mathbf{v})$. However, for each $\preceq \in \mathscr{B}'$, there may be multiple consistent valuations in $\mathscr{V}$.

For each player i, let $\preceq_i$ denote the element of $\mathscr{B}'(\mathbf{v})$ that is consistent with her true values v_i. After nature draws v_i, nature either reveals it to player i with probability $\alpha \in [0,1]$ or nature reveals $\preceq_i$ with probability $1-\alpha$; the lottery is independent of all other randomness. If i observes v_i, then she knows her cardinal preferences; however, if she observes $\preceq_i$, then she only knows her ordinal preferences.

Let θ_i denote nature's report to player i, which we refer to i's **type**. Let $\Theta = \mathscr{V} \cup \mathscr{B}'$ denote the set of types. The probability i gets type θ_i is,[10]

$$\Pr(\theta_i) = \begin{cases} \alpha f_v(\theta_i) & \text{if } \theta_i \in \mathscr{V} \\ (1-\alpha)\sum_{v \in I(\theta_i)} f_v(v) & \text{if } \theta_i \in \mathscr{B}', \end{cases} \tag{1}$$

where $I(\theta_i) = \{v \in \mathscr{V} | v \text{ is consistent with } \theta_i\}$. Let $\theta = (\theta_1, \ldots, \theta_N)$ denote the profile of players' types. Due to independence, the probability of θ is $\Pr(\theta) = \Pi_{i=1}^N \Pr(\theta_i)$. When $\alpha \in (0,1)$, all type profiles occur with strictly positive probability because $f_v > 0$.

After each player observes her type, she makes a report about her preferences to an assignment mechanism, which processes the reports and generates a (mixed) assignment; we focus on direct mechanisms for simplicity. Let $r_i \in \Theta$ denote the report of player i and $r = (r_1, \ldots, r_N)$ denote the profile of players' reports. Formally, a **mechanism** is a $M : \Theta^N \to \Delta\Phi$. After the mechanism generates its mixed assignment, the outcome of the assignment is realized and players receive their objects and their corresponding true values.

The players are Bayesians who have Von-Neumann Morgenstern utility. Thus, when player i gets type θ_i and makes her report, her expected true value

[10] If $\theta_i \in \mathscr{V}$, then nature must have drawn θ_i for player i and revealed it to her, an event of probability $\alpha f_v(\theta_i)$. If $\theta_i \in \mathscr{B}'$, then nature must have drawn a vector v of true values that are consistent with θ_i and must not have revealed v to i, an event of probability $(1-\alpha)\sum_{v \in I(\theta_i)} f_v(v)$.

vector is,[11]

$$v_i^\dagger = (v_{\eta i}^\dagger, v_{1i}^\dagger, \ldots, v_{oi}^\dagger, \ldots, v_{Ki}^\dagger) = \begin{cases} \theta_i & \text{if } \theta_i \in \mathscr{V} \\ \sum_{v \in I(\theta_i)} v \frac{f_v(v)}{\sum_{v' \in I(\theta_i)} f_v(v')} & \text{if } \theta_i \in \mathscr{B}'. \end{cases} \tag{2}$$

And her payoff, denoted $u_i(\cdot|\theta_i)$, is,

$$u_i(o|\theta_i) = v_{oi}^\dagger \text{ for each } o \in \mathscr{O}'.$$

In abuses of notation, (i) we write $u_i(\phi|\theta_i) = u_i(\phi(i)|\theta_i)$ for i's payoff to the pure assignment ϕ and (ii) we write $u_i(\tilde{\phi}|\theta_i) = \sum_{j=1}^{|\Phi|} u_i(\phi_j|\theta_i)\tilde{\phi}_j$ for i's payoff to in the mixed assignment $\tilde{\phi} = (\tilde{\phi}_1, \ldots, \tilde{\phi}_{|\Phi|})$.

EFFICIENCY AND TRUTHFUL REPORTING

We introduce three definitions to formalize our efficiency and incentive compatibility notions. Our first pair of definitions concerns the efficiency of assignments and mechanisms.

Given a profile of players' types $\theta = (\theta_1, \ldots, \theta_N) \in \Theta^N$, we say that the (mixed) assignment $\tilde{\phi}^\star$ is **interim efficient** if it solves $\max_{\tilde{\phi} \in \Delta\Phi} \sum_{i \in \mathscr{N}} u_i(\tilde{\phi}|\theta_i)$, i.e., if it maximizes utilitarian welfare after nature has endowed types but before all uncertainty about true values has been resolved. If $\tilde{\phi}^\star$ is not interim efficient, we say it is **interim inefficient**.

An interim efficient (mixed) assignment $\tilde{\phi}^\star$ is also "interim" ("ex-post") Pareto efficient, i.e., there is no other pure assignment that leaves all players weakly better off and one player strictly better off before (after) all uncertainty about true values has been resolved.[12] Interim and ex-post Pareto efficiency imply that $\tilde{\phi}^\star$ is also "individually rational," i.e., each player gets an element of $\mathscr{O}'$ she likes at least as much as η both before and after all uncertainty about true values is resolved.[13] In addition, when $\mathscr{V}$ contains only "strict value

[11] There are only two types of zero probability events, (i) $\alpha = 0$ and i observes a type $\theta_i \in \mathscr{V}$, and (ii) $\alpha = 1$ and i observes a type $\theta_i \in \mathscr{B}'$. If (i), then we assume that θ is actually i's vector of true values. If (ii), then we assume that i's true value is drawn from $I(\theta)$ with probability $\frac{f_v(v)}{\sum_{v' \in I(\theta)} f_v(v')}$ for each v therein. The display equation embeds both of these natural assumptions.

[12] It is evident that interim efficient implies interim Pareto efficiency; we establish that it also implies ex-post Pareto efficiency in the Online Appendix.

[13] The individual rationality criterion reflects the idea that each player does weakly better by participating in the assignment instead of opting out and getting η with certainty.

vectors,[14]" then $\tilde{\phi}^{\star}$ is "ordinally efficient" by Bogomolnaia & Moulin (2001) Lemma 2.

Since payoffs are continuous in the mixed assignment and the set of mixed assignments is compact, the Extreme Value Theorem applies and guarantees the existence of an interim efficient assignment. We have thus proved the following result.

Lemma 2. Existence of an Interim Efficient Assignment.
For each $\theta \in \Theta^N$, there is an interim efficient mixed assignment $\tilde{\phi}^{\star}$.

We say that a mechanism M is **interim efficient** if $M(\theta)$ is an interim efficient (mixed) assignment for each profile of types $\theta \in \Theta^N$. In other words, an interim efficient mechanism is one that always selects an interim efficient assignment when the players truthfully report their types. Such a mechanism inherits the interim and ex-post Pareto efficiency of interim efficient assignments, as well as their ordinal efficiency and individual rationality. If M not interim efficient, we say that M is **interim inefficient**.

Our third definition concerns players' incentives to truthfully report their types to a mechanism. For player i, let $U_i^M(r_i|\theta_i)$ denote i's payoff to reporting $r_i \in \Theta$ to the mechanism M given her type θ_i, when all other players truthfully report their types $\theta_{-i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_N)$, i.e.,

$$U_i^M(r_i|\theta_i) = \sum_{\theta_{-i} \in \Theta^{N-1}} u_i(M(r_i, \theta_{-i})|\theta_i) \Pr(\theta_{-i}),$$

where $\Pr(\theta_{-i}) = \Pi_{j \neq i} \Pr(\theta_j)$ is the probability of θ_{-i}.

For each $\varepsilon > 0$, a mechanism M is **ε-Bayesian incentive compatible** if, for each player i and each type $\theta_i \in \Theta$, i does not gain more than ε from lying strategically when everyone else tells the truth, i.e., if $\max_{r_i \in \Theta} U_i^M(r_i|\theta_i) \leq U_i^M(\theta_i|\theta_i) + \varepsilon$.[15] Since it is often costly – in terms of effort and time – for a person to formulate a strategic manipulation, ε-Bayesian incentive compatibility is a compelling truth-telling notion when ε is small – as then the gains

[14] That is, $(v_\eta, v_1, \ldots, v_K) \in \mathscr{V}$ implies $v_o \neq v_{o'}$ for any two distinct elements o and o' of $\mathscr{O}'$.

[15] The ε-Bayesian incentive compatibility truth-telling notion is weaker than other classical truth-telling notions, including: (i) "strategy-proofness," which requires players find truth-telling optimal when they know each other's types, and (ii) "Bayesian incentive compatibility," which requires that ε-Bayesian incentive compatibility hold for $\varepsilon = 0$. It is, however, similar in spirit to other "approximate" truth-telling notions such as "strategy-proofness in the large" (Azevedo & Budish, 2018) and "asymptotic strategy-proofness" (Liu & Pycia, 2016).

Player		"Objects" η	o	o'
1	θ_1	1	3	2
2	θ_2	1	5	3
3	θ_3	2nd	1st	3rd

(a) Initial Types

Probability	Player 1	2	3
1/3	o	o'	η
1/3	o'	o	η
1/6	o'	η	o
1/6	η	o'	o

(b) Deferred Acceptance with Uniform Tie-Breaking

Player		"Objects" η	o	o'
1	θ_1'	1	5	3
2	θ_2'	1	5	3
3	θ_3'	2nd	1st	3rd

(c) New Types

Table 1: Illustration

from manipulation are less than the (implicit) costs of manipulation. Hence, truth-telling is a Nash equilibrium of the reporting phase.

3. ILLUSTRATION

To illustrate our approach and give intuition to our core results, consider a simple example with three players, labeled 1 to 3, and two objects, labeled o and o'.[16] Let the chance nature reveals a true value be $\alpha = \frac{1}{2}$, let $\mathscr{V}$ have five elements – $(v_\eta, v_o, v_{o'}) = (1,3,2)$, $(v_\eta, v_o, v_{o'}) = (1,5,3)$, $(v_\eta, v_o, v_{o'}) = (2,3,0)$, $(v_\eta, v_o, v_{o'}) = (3,4,1)$, and $(v_\eta, v_o, v_{o'}) = (1,6,7)$ – and let each element be equally likely, i.e., $f_v = \frac{1}{5}$.

TYPE, PAYOFFS, AND AN INTERIM EFFICIENT ASSIGNMENT

Suppose the players' types are given by Table 3.1(a). So, players 1 and 2 learn their cardinal preferences and thus their payoffs, which are $u_1(y) = 1\mathbb{I}(y = \eta) + 3\mathbb{I}(y = o) + 2\mathbb{I}(y = o')$ and $u_2(y) = 1\mathbb{I}(y = \eta) + 5\mathbb{I}(y = o) + 3\mathbb{I}(y = o')$ respectively.[17] Player 3 only learns her ordinal preference. Since $(v_\eta, v_o, v_{o'}) = (2,3,0)$ and $(v_\eta, v_o, v_{o'}) = (3,4,1)$ are the only true values consistent with 3's type and since all types are equally likely, her expected true values are

$$\frac{\frac{1}{5}}{\frac{2}{5}}(2,3,0) + \frac{\frac{1}{5}}{\frac{2}{5}}(3,4,1) = (\frac{5}{2}, \frac{7}{2}, \frac{1}{2}),$$

and her payoff is $u_3(y) = \frac{5}{2}\mathbb{I}(y = \eta) + \frac{7}{2}\mathbb{I}(y = o) + \frac{1}{2}\mathbb{I}(y = o')$.

[16] We refer to the objects as o and o' to avoid confusion with player names.

[17] We write $\mathbb{I}(\{\text{event}\})$ for the classic indicator function, which equals one when event is true and equals zero otherwise. In addition, we suppress the dependency of payoffs on type for expositional simplicity.

In the corresponding interim efficient assignment, player 1 gets o', player 2 gets o, and player 3 gets η (i.e., 3 is left unassigned). Simply, its surplus maximizing to give objects to those who value them most, so: (i) the fact player 2 has the highest value for o (among all the players) implies it is efficient for her to get it and, given this, (ii) the facts players 1 and 3 have highest values for o' and η respectively (among all the remaining players) imply it is efficient they be assigned these items.

RECOMMENDATIONS OF CLASSIC MECHANISMS

Many classic mechanisms do not suggest the interim efficient assignment (when players report truthfully). For instance, Deferred Acceptance with Uniform Tie-Breaking, as well as Random Serial Dictatorship and Top Trading Cycles with Random Endowments, yield the mixed assignment in Table 3.1(b).[18] This assignment is interim inefficient: its social surplus is $\frac{25}{3} \approx 8.33$, while the social surplus of the interim efficient assignment is $\frac{19}{2} = 9.5$, for an inefficiency of $\frac{7}{6}$.[19]

Intuitively, these mechanisms do not generate the interim efficient assignment because they "disregard cardinal information" and thus do not award objects to the players who value them most. To illustrate, consider Deferred Acceptance with Uniform Tie-Breaking. This mechanism consumes players reports and then, on behalf of each player, it "proposes" to the player's most-preferred element in $\mathscr{O}'$. Each object then accepts one of its proposals, breaking ties at random; the object is then assigned to the player whose proposal it accepted. (Players who propose to η are always assigned to it.) Afterwards, the mechanism re-proposes on behalf of each player whose proposal was rejected

[18] Deferred Acceptance traces its origins back to the classical "marriage market" of Gale & Shapley (1962), and finds more recent treatments in Abdulkadiroglu et al. (2011) and Roth & Sotomayor (1990); we describe it below. The Random Serial Dictatorship mechanism randomly orders players according to a uniform distribution, and then assigns the first her most-preferred element in $\mathscr{O}'$, the second her most preferred element among the remaining elements, and so on. Abdulkadiroglu & Sonmez (1998) establish that the mixed outcome of Random Serial Dictatorship is the same as under Top Trading Cycles with Random Endowments. This latter mechanism endows players randomly with objects or η and asks them to "point" to their most preferred objects, a cycle of the implied graph is then found and its trades implemented as assignments; the process repeats until there are no cycles left. Top Trading Cycles traces its origins back to Shapley & Scarf (1974), and finds modern treatments in Abdulkadiroglu & Sonmez (2003) and Roth et al. (2004).

[19] While this inefficiency is small, inefficiencies in general can be quite large – e.g., if we multiply the true values by $\alpha > 1$, then the inefficiency is $\frac{7}{6}\alpha$.

to one of the remaining objects or η. The process concludes once all players are assigned to an object or η.

Since the three players initially propose to object o, there is a $\frac{1}{3}$ chance each gets it due to o's uniform tie-breaking. Thus, the mixed assignment's outcome is inefficient $2/3$-rds of the time.[20] The root of the inefficiency lies in the fact that o ignores players' expected true values when deciding which proposal to accept. Instead, if o broke ties in favor of higher values for itself, then (i) it would go to player 2 in the initial round and, in the second round, (ii) o' would go to player 1 and η would go to player 3. Thus, the interim efficient assignment would obtain with certainty.

This leads us to a classical conclusion of the random assignment literature: mechanisms that disregard cardinal information "can be" interim inefficient. In Proposition 1, we layout broad conditions which strengthen this conclusion and transform "can be" into "must be." Roughly, these conditions stipulate that (i) there are a sufficient number of players and (ii) that $\mathscr{V}$ contains two different, possible vectors of true values that induce the same ordinal preference. We thus obtain the following observation.

Observation 1. Mechanisms that disregard cardinal information are interim inefficient.

One might think that the inefficiency problem is limited only to mechanisms that disregard cardinal information. In Section 4, we show that this is not the case by establishing that classic mechanisms that use cardinal information, like Pseudo-markets, are also interim inefficient. The broad intuition is that mechanisms often use cardinal information in ways that are unrelated to social surplus maximization.

THE SIMPLE MECHANISM

Motivated by Observation 1, we introduce the "Simple Mechanism," which is formally described in Section 5. In this mechanism, players submit either cardinal or ordinal reports. The mechanism uses these reports to construct an

[20] The mixed assignment in Table 3.1(b) results from this tie-breaking. If player 1 wins o, then player 2 proposes to o' and player 3 proposes to η in the next round. Since both proposals are accepted, the assignment in the first row of Table 3.1(b) is obtained. If 2 wins o, then 1 proposes to o' and 3 proposes to η in the next round, so the assignment in the second row of Table 3.1(b) is obtained. If 3 wins o, then 1 and 2 both propose to o' in the next round, with the loser proposing to η in the third round, so the assignments in the last two rows of Table 3.1(b) are obtained. The probabilities of each pure assignment in Table 3.1(b) are product of the respective tie-breaking events.

estimate of each player's payoffs. It does this by treating a player's report as her type and computing her expected true values: specifically, (i) if a player submits a cardinal report r, then the mechanism estimates her expected true values are r, and if a player submits an ordinal report r, then the mechanism estimates her expected true values are $\sum_{v \in I(r)} (v f_v(v) / \sum_{v' \in I(r)} f_v(v'))$. The mechanism then computes the set of pure assignments σ that maximizes the sum of these estimated payoffs (i.e., maximize its estimate of social surplus) and it selects one of these assignments to implement with uniform probability.

To illustrate how the Simple Mechanism works, suppose the players truthfully report their types from Table 3.1(a). Then the mechanism estimates players 1 and 2 have payoffs of $\hat{u}_1(y) = 1\mathbb{I}(y = \eta) + 3\mathbb{I}(y = o) + 2\mathbb{I}(y = o')$ and $\hat{u}_2(y) = 1\mathbb{I}(y = \eta) + 5\mathbb{I}(y = o) + 3\mathbb{I}(y = o')$ respectively since they make cardinal reports. Since player 3 submits an ordinal report r_3, the mechanism first determines that $I(r_3) = \{(2,3,0),(3,4,1)\}$ and estimates her expected true values are $(\frac{5}{2}, \frac{7}{2}, \frac{1}{2}) = \sum_{v \in I(\theta_3)} (v \frac{1}{5} / \sum_{v' \in I(\theta_3)} \frac{1}{5})$ and her payoff is $\hat{u}_3(y) = \frac{5}{2}\mathbb{I}(y = \eta) + \frac{7}{2}\mathbb{I}(y = o) + \frac{1}{2}\mathbb{I}(y = o')$. The mechanism then solves $\max_{\phi \in \Phi} \sum_{i \in \{1,2,3\}} \hat{u}_i(\phi(i))$ and concludes that the unique solution is for player 1 to get o', player 2 to get o, and player 3 to get η. And it implements this solution with certainty.

The interim efficiency of this solution is by design: when players tell the truth, the Simple Mechanism correctly estimates their payoffs and thus maximizes social welfare when it maximizes the sum of their estimated payoffs. This leads to our second observation, which we formalize in Proposition 2.

Observation 2. The Simple Mechanism is interim efficient.

It follows that the Simple Mechanism is also interim and ex-post Pareto efficient, as well as individually rational. Nevertheless, the Simple Mechanism may be inefficient when players lie. To see this, suppose player 3 misreports her preference and states that o' is most preferred, o is second most preferred, and η is least preferred. Then, the mechanism estimates her payoff is $\hat{u}'_3(y) = 1\mathbb{I}(y = \eta) + 6\mathbb{I}(y = o) + 7\mathbb{I}(y = o')$, maximizes $\hat{u}_1 + \hat{u}_2 + \hat{u}'_3$, and concludes that there is a unique solution: assign η to player 1, assign o to player 2, and assign o' to player 3. This solution is clearly interim inefficient.

A positive feature of the Simple Mechanism is that it is "symmetric," i.e., it treats equals equally. To illustrate, suppose Table 3.1(c) gives the players' types, instead of Table 3.1(a); so, players 1 and 2 now have the same cardinal preference. Then, under truthful reporting, the Simple Mechanism computes

that there are two assignments in σ: (i) 1 gets o, 2 gets o', and 3 gets η and (ii) 1 gets o', 2 gets o, and 3 gets η. Since it mixes over these assignments with equal probability, players 1 and 2 both have a $\frac{1}{2}$ chance of getting o or o' and thus have the same payoff.

The intuition straightforward. Since players 1 and 2 have the same payoff function, it is always possible to have them exchange objects while retaining efficiency. Thus, for every assignment in σ where player 1 does better than player 2 by an amount δ, there is a second assignment in σ where 2 does better than 1 by δ. Since the Simple Mechanism randomizes uniformly over σ, both players have the chance of getting the better or worse assignment and thus have the same payoff. This leads us to our third observation, which we formalize in Proposition 2.

Observation 3. When two players have the same preferences and report truthfully, then they have the same payoff to the Simple Mechanism.

There is a question of whether players want to tell the truth to the Simple Mechanism. Unfortunately, truth-telling is neither strategy-proof nor Bayesian incentive compatible. To illustrate the lack of Bayesian incentive compatibility (and thus the lack of strategy-proofness),[21] it helps to simplify the example by assuming that there are only two players, 1 and 2, that $\alpha = 1$, and that $\mathscr{V}$ contains two equally likely elements $(v_\eta, v_o, v_{o'}) = (1,5,3)$ and $(v_\eta, v_o, v_{o'}) = (1,3,2)$, which we label A and B respectively. Focus on player 1 and suppose her type is B.

First, consider player 1's payoff to truth-telling, i.e., reporting she is a type B, when player 2 tells the truth. If player 2 is a type A, then the Simple Mechanism gives 2 object o and gives 1 object o'. If player 2 is type B, then the Simple Mechanism computes there are two assignments in σ: (i) 1 gets o and 2 gets o' and (ii) 1 gets o' and 2 gets o; so, player 1 gets o half of the time and gets o' the other half of the time. Hence, player 1's payoff to truth-telling is $\frac{1}{2}2 + \frac{1}{2}(\frac{1}{2}3 + \frac{1}{2}2) = \frac{9}{4}$.

Second, consider player 1's payoff to lying, i.e., reporting she is a type A, when player 2 tells the truth. If player 2 is a type A, then the Simple Mechanism computes that there are two assignments in σ: (i) 1 gets o and 2 gets o' and (ii) 1 gets o' and 2 gets o; so, player 1 gets o half of the time and gets o' the other

[21] Strategy-proofness implies Bayesian incentive compatibility because the former requires truth-telling be optimal for each possible profile of types, whereas the latter integrates across type profiles.

half of the time. If player 2 is type B, then the Simple Mechanism concludes that it is best to give o to 1 and give o' to 2. Hence, player 1's payoff to lying is $\frac{1}{2}(\frac{1}{2}3+\frac{1}{2}2)+\frac{1}{2}3=\frac{11}{4}$.

Player 1's gain to lying is thus $\frac{1}{2}=\frac{11}{4}-\frac{9}{4}$. Intuitively, this gain is positive because the Simple Mechanism awards objects to those who it thinks value them most. Thus, if player 1 does not report the highest value for o', then she will not get it with positive probability. Since player 1 knows this and knows she has the lowest value for o', she finds it best to lie and misreport her type.[22]

Importantly, player 1's gain to lying decreases as the number of truth-telling players $N-1$ grows. To see this, let A_l be the event that there are l players (aside from player 1), who report type A, i.e., $A_l=\{\theta_{-1}|\theta_{-1}$ specifies l players have type $A\}$. We can then write 1's payoff to reporting r_1 as

$$U_1^{M_S}(r_1|B)=\sum_{l=0}^{N-1}\sum_{\theta_{-1}\in A_l}u_1(M_S(r_1,\theta_{-1})|B)\Pr(\theta_{-1}),$$

where M_S denotes random assignment selected by the Simple Mechanism. It is straightforward to compute the value of $u_1(M_S(r_1,\theta_{-1})|B)$ on each set A_l:

- For $l=0$ and all θ_{-1} in A_l, we have

$$u_1(M_S(r_1,\theta_{-1})|B)=\begin{cases}3 & \text{if } r_1=A\\ \frac{3}{N}+\frac{2}{N}+1(1-\frac{2}{N})=1+\frac{3}{N} & \text{if } r_1=B.\end{cases}$$

  Since type As value both o and o' strictly more than type Bs, the Simple Mechanism gives o and o' to players who report type A in order to maximize its estimate of social surplus. Thus, (i) if player 1 reports A, then she gets o since there are no other type A players, and (ii) if she reports B, then she gets o' with probability $\frac{1}{N}$, o with probability $\frac{1}{N}$, and η with probability $(1-\frac{2}{N})$.[23]

[22] Ehlers et al. (2014) leverage an analogous intuition to show that every mechanism that awards objects based on cardinal information is not Bayesian incentive compatible in finite assignment environments.

[23] These probabilities come from a counting exercise. There are $\binom{N}{1}\binom{N-1}{1}$ ways to distribute the two objects to the N players who report B. Thus, if player 1 has o (or o'), there are $\binom{N-1}{1}$ ways to distribute o' (o) to the other players. Since all assignments in σ are equally likely, the probability player 1 gets o (o') is $\binom{N-1}{1}/(\binom{N}{1}\binom{N-1}{1})=\frac{1}{N}$. Since player 1 gets η if she does not get o or o', the probability of η is $1-\frac{2}{N}$. (The counting exercises for all subsequent bullets are analogous and omitted.)

- For $l = 1$ and all θ_{-1} in A_l, we have

$$u_1(M_S(r_1,\theta_{-1})|B) = \begin{cases} \frac{3}{2}+\frac{2}{2}=\frac{5}{2} & \text{if } r_1 = A \\ \frac{2}{N-1}+1(1-\frac{1}{N-1}) = 1+\frac{1}{N-1} & \text{if } r_1 = B. \end{cases}$$

  If $r_1 = A$, then the Simple Mechanism assigns both objects to the two A reporters, implying that player 1 has an equal chance of getting either object. If $r_1 = B$, then because the player who truthfully reports A always gets o, the chance player 1 gets o' is $\frac{1}{N-1}$.

- For $l \geq 2$ and all θ_{-1} in A_l, we have

$$u_1(M_S(r_1,\theta_{-1})|B) = \begin{cases} \frac{1}{l+1}3+\frac{1}{l+1}2+(1-\frac{2}{l+1})1 = 1+\frac{1}{l+1}3 & \text{if } r_1 = A \\ 1 & \text{if } r_1 = B. \end{cases}$$

  If $r_1 = A$, then player 1 gets o with probability $\frac{1}{l+1}$, o' with probability $\frac{1}{l+1}$, and η with probability $(1-\frac{2}{l+1})$. If $r_1 = B$, then player 1 gets η because o and o' go to the A reporters.

Hence,

$$U_1^{M_S}(r_1|B) = \begin{cases} 3\Pr(A_0)+\frac{5}{2}\Pr(A_1)+\sum_{l=2}^{N-1}(1+\frac{3}{l+1})\Pr(A_l) & \text{if } r_1 = A \\ (1+\frac{3}{N})\Pr(A_0)+(1+\frac{1}{N-1})\Pr(A_1)+\sum_{l=2}^{N-1}\Pr(A_l) & \text{if } r_1 = B, \end{cases} \tag{3}$$

where, due to the independence of players' types, $\Pr(A_l) = \sum_{\theta_{-1}\in A_l}\Pr(\theta_{-1}) = \binom{N-1}{l}\frac{1}{2}^{N-1}$.[24] Observe that $U_1^{M_S}(r_1|B)$ is greater than one and decreases to one as N grows large.[25] Thus, for each $\varepsilon > 0$, there is a finite $\bar{N} > 0$ such that $1 < U_1^{M_S}(A|B) \leq 1+\varepsilon$ and $1 < U_1^{M_S}(B|B) \leq 1+\varepsilon$ when there are at least $\bar{N}$

[24] There are $\binom{N-1}{l}$ elements in A_l and the probability of each is $\left(\frac{1}{2}\right)^{N-1}$ since types are independent, $\alpha = 1$, and all both types are equally likely.

[25] We establish that $\lim_{N\to\infty} U_1^{M_S}(A|B) = 1$ since the argument that $\lim_{N\to\infty} U_1^{M_S}(B|B) = 1$ is analogous. Since $\lim_{N\to\infty}\Pr(A_l) = 0$, it is evident that the first two terms on the right-hand-side of equation (3) go to zero. We use a squeeze argument to show that the last term goes to one. Let $\delta > 0$. Since $\frac{3}{l+1}$ is strictly decreasing, there is a fixed N' such that (i) $\frac{3}{l+1} > \delta$ for all $l < N'$ and (ii) $\frac{3}{l+1} \leq \delta$ for all $l \geq N'$. Thus, for large N, we have $1 \leq \sum_{l=2}^{N-1}(1+\frac{3}{l+1})\Pr(A_l) \leq \sum_{l=2}^{N'-1}(1+\frac{3}{l+1})\Pr(A_l)+(1+\delta)\sum_{l=N'}^{N-1}\Pr(A_l)$. Since $\sum_{l=N'}^{N-1}\Pr(A_l) = (1-\sum_{l=0}^{N'-1}\Pr(A_l))$, taking limits yields $1 \leq \lim_{N\to\infty}\sum_{l=2}^{N-1}(1+\frac{3}{l+1})\Pr(A_l) \leq 1+\delta$. Letting $\delta \to 0$ gives the desired result.

players. In other words, player 1 never gains more than ε by lying when there are at least $\bar{N}$ players.[26]

The economic intuition is simple: "competition" for the objects increases as the number of truthful type A players rises (due to the growth of N). This competition is driven by the Simple Mechanism's randomization, which ensures that each A reporter has an equal chance of getting either object. Thus, the probability that player 1 can obtain either object by lying or telling the truth falls, limiting her ability to manipulate the mechanism. Since player 1 is symmetric to all type B players and since no type A player has incentive to lie, we obtain the following observation, which we formalize in Proposition 3.

Observation 4. For every $\varepsilon > 0$, the Simple Mechanism is ε-Bayesian incentive compatible for large N.

The Simple Mechanism thus has good truth-telling properties when the number of players is large. This, along with interim efficiency and symmetry, suggest that the Simple Mechanism may be a good mechanism to employ in situations there is substantial demand for the objects – as might be the case in the assignment of prime university housing or courses, in the awarding of grants or fellowships, in the allocation of slots in programs for free/subsidized day-care or housing, in the apportionment of prime employee parking spots or offices, or even in the distribution of kidneys and other organs.

We have held the set of objects fixed while increasing the number of players. While there are situations where the supply of objects is fixed, it is often natural to think that it increases with the number of players – e.g., colleges build dorm rooms and add courses in response to rising enrollments and local governments expand free day-care in response to growing populations. Fortunately, this assumption is for expositional simplicity. As we discuss in Proposition 4, the set of objects may increase with the number of players via replication, provided it does so slowly enough to maintain competitive pressure and thus approximate incentive compatibility.

[26] In practice, N does not have to be very large for the gain to lying to be small – e.g., computation gives that player 1's gain to lying is 0.12 when there are 50 players and 0.06 when there are 100 players.

4. CARDINAL INFORMATION AND INEFFICIENCY

Intuitively, mechanisms that "disregard cardinal information" can be interim inefficient. Our main goal, in this section, is to strengthen this intuition and show that such mechanisms "must be" interim inefficient under broad conditions. We also discuss the efficiency of classical mechanisms that make use of cardinal information.

Disregarding Cardinal Information

We have in mind that a mechanism "disregards cardinal information" if it only uses the ordinal information in players' reports. To make this notion precise, for each $\theta \in \Theta$, let $c(\theta)$ give the order that is consistent with θ. Thus, (i) $c(\theta) = \theta$ if $\theta \in \mathscr{B}'$ or (ii) $c(\theta) = \preceq$ if $\theta \in \mathscr{V}$ where $\preceq$ is the element of $\mathscr{B}'$ that is consistent with θ. For a type profile $\theta = (\theta_1, \ldots, \theta_N)$ or a report profile $r = (r_1, \ldots, r_N)$, we write $c(\theta)$ and $c(r)$ for $(c(\theta_1), \ldots, c(\theta_N))$ and $(c(r_1), \ldots, c(r_N))$ respectively for players ordinal preferences and ordinal reports. We say that a mechanism M **disregards cardinal information** if $M(r) = M(c(r))$ for all $r \in \Theta^N$. Many classic mechanisms disregard cardinal information since they only use players' rankings of objects when they run – e.g., Deferred Acceptance, Probabilistic Serial,[27] Random Serial Dictatorship, and Top Trading Cycles.

To establish the interim inefficiency of mechanisms that disregard cardinal information, we place a regularity condition on $\mathscr{O}$ and $\mathscr{V}$.

Assumption 1. There exist two objects o and o' in $\mathscr{O}$ and two vectors $\check{v} = (\check{v}_\eta, \ldots, \check{v}_o, \ldots, \check{v}_{o'}, \ldots, \check{v}_K)$ and $\check{v}' = (\check{v}'_\eta, \ldots, \check{v}'_o, \ldots, \check{v}'_{o'}, \ldots, \check{v}'_K)$ in $\mathscr{V}$ such that:
(i) Both vectors induce the same ordinal preferences, i.e., $c(\check{v}) = c(\check{v}')$.
(ii) Object o' is strictly preferred to o and o is strictly preferred to η, i.e., $\check{v}_{o'} > \check{v}_o > \check{v}_\eta$ and $\check{v}'_{o'} > \check{v}'_o > \check{v}'_\eta$.
(iii) The "marginal" value to giving up o to get o' is larger in $\check{v}'$ than $\check{v}$, i.e., $\check{v}'_{o'} - \check{v}'_o > \check{v}_{o'} - \check{v}_o$.

Assumption 1 is a mild condition that is satisfied, for instance, in the Illustration or, more generally, when $\mathscr{V} = \{v_0, v_1, \ldots, v_L\}^{K+1}$ with $v_0 < v_1 < \cdots < v_L$,

[27] The Probabilistic Serial mechanism was introduced by Bogomolnaia & Moulin (2001). In this mechanism, players consume their most preferred object at a some "eating speed." Once the objects are consumed, players move onto their second most preferred objects, provided they are available, and so on until all objects are consumed. The proportion of each object eaten by player i then corresponds to the probability the that i gets the object in the assignment.

$K \geq 2$, and $L \geq 3$.[28]

Proposition 1. Disregarding Cardinal Information and Inefficiency.
Let Assumption 1 *hold, let* $\alpha > 0$*, and let* M *be a mechanism that disregards cardinal information. Then, there is a finite* $\bar{N} > 0$ *such that* M *is interim inefficient when there are at least* $\bar{N}$ *players.*

The proposition implies, for instance, that Deferred Acceptance, Probabilistic Serial, Random Serial Dictatorship, and Top Trading Cycles all produce assignments with social surplus that is *strictly* less than what could be achieved for at least one type profile. The size of the inefficiencies may be economically substantial, especially when the values in $\mathscr{V}$ are large or "spread out."

The intuition for the result mirrors the intuition behind Observation 1; we exploit it in the proof by initially endowing all players with the same type and selecting a player who gets o under M. We then increase this player's value to o' so that it becomes efficient for her alone to get it. Since this cardinal perturbation leaves ordinal types unaltered, M still assigns o to our player and is thus interim inefficient. Assumption 1 enables this preference alteration.

The lower bound on the number of players rules out corner cases where interim efficiency may obtain even for mechanisms that disregard cardinal information – e.g., when there is a single player, Randomized Serial Dictatorship always allows her to obtain her most preferred object, an interim efficient outcome. The restriction that $\alpha > 0$ ensures that some players have cardinal preferences; if $\alpha = 0$, then all preferences are ordinal and the proposition is vacuous.

Making Use of Cardinal Information

In light of Proposition 1, it is natural to wonder if mechanisms that "use cardinal information" are interim efficient. The answer, unfortunately, is "not always:" some, like the Simple Mechanism, are efficient, while other classic mechanisms, like "Nash-Boston" or Pseudo-markets, are not.[29]

Nash-Boston Mechanism. While the Boston mechanism only solicits players' strict ordinal preferences and so is interim inefficient, the equilibrium of its reporting game depends on players best responses, which themselves

[28] When $K \geq 2$ and $L \geq 3$, there are value vectors $(v_0, v_0, \ldots, v_0, v_1, v_2)$ and $(v_0, v_0, \ldots, v_0, v_1, v_3)$, which are easily seen to meet (i), (ii), and (iii) of Assumption 1 because $v_3 > v_2 > v_1$.

[29] Interestingly, when there is a continuum of players and a mass of objects, both Nash-Boston and Pseudo-markets are interim efficient for a broad class of environments – see Hafalir & Miralles (2014).

depend on their preference intensities; see, for instance, Abdulkadiroglu et al. (2011) and Miralles (2008). It is thus natural to ask whether this Nash equilibrium – which we refer to as the "Nash-Boston" mechanism since it maps types into allocations – constitutes an interim efficient mechanism. As the next example shows, the answer is no.

Example 1. Nash-Boston and Interim Inefficiency.

Let there be two players, 1 and 2, and two objects, o and o'. Let $\alpha = 1$ and let $\mathscr{V}$ contain two equally likely elements $(v_\eta, v_o, v_{o'}) = (0, 10, -1)$ and $(v_\eta, v_o, v_{o'}) = (0, 9, 8)$, which we label A and B respectively. It is clear that if both players are type A (B), then the two interim efficient pure assignments give o to one player and η to the other player (give o to one player and o' to the other player). However, if there is one of each type, then the interim efficient assignment gives o to the type A player and o' to the type B player.

Computation shows that, in the (unique) symmetric equilibrium of the Boston Mechanism's reporting game, both types of players truthfully report their ordinal preferences.[30] This equilibrium has the following outcomes. If both players are type A (B), one is randomly assigned o and the other is assigned η (assigned o and the other is assigned o'). If one player is type A and the other is type B, one is randomly assigned o and the other is assigned either η if type A or o' if type B. In all outcomes, both players have an equal probability of being assigned o.

The equilibrium type-assignment mapping, i.e., the Nash Boston mechanism, produces a mixed assignment that randomizes over the two interim efficient pure assignments when both players are type A (B) and so itself is interim efficient. However, when one player is type A and the other type B, then the mechanism produces a mixed assignment that gives o to the type B player with non-zero probability and so is interim inefficient. It follows that the Nash-Boston mechanism is interim inefficient: intuitively, while the equilibrium consumes preference intensities via players' best responses, this information is used to form a mutual best response and not to cooperatively maximize social surplus. $\triangle$

Pseudo-Market Mechanisms. A Pseudo-market mechanism assigns goods according to the Walrasian equilibrium of an artificial economy that is constructed from players' reported preferences; see, for instance, Hylland &

[30] We defer the details of the Boston mechanism, its reporting game, and the equilibrium to the Appendix for expositional simplicity.

Zeckhauser (1979).[31] Specifically, the mechanism endows each player with an artificial budget, which is specified in the mechanism's description, and then maximizes her reported utility by buying probability shares of the objects subject to this constraint, while taking prices as given. It then solves for the Walrasian equilibrium and uses the equilibrium probability shares as the basis for the assignment. Since the Pseudo-market uses cardinal information in the purchase of probability shares, one might think that its equilibrium is interim efficient. The next example shows that this is not the case.

Example 2. A Pseudo-Market and Interim Inefficiency.

Let there be three players, 1, 2, and 3, and two objects, o and o'. Let $\alpha = 1$ and let $\mathscr{V}$ contains three equally likely elements $(v_\eta, v_o, v_{o'}) = (0,4,2)$, $(v_\eta, v_o, v_{o'}) = (0,3,2)$, and $(v_\eta, v_o, v_{o'}) = (0,2,4)$, which we label A, B, and C respectively.

To begin, suppose that nature draws type A for player 1, type B for player 2, and type C for player 3, i.e., types are $\theta = (A,B,C)$. Then, it is uniquely interim efficient for o go to player 1, for η to go to player 2, and for o' go to player 3.

Consider a Pseudo-market mechanism M, where (i) players 1 and 3 have budgets of one and (ii) player 2 has a budget of zero. Suppose the prices for o and o', denoted p_o and $p_{o'}$, are $p_o = p_{o'} = 1$. On behalf of player 1, the mechanism solves $\max_{(x_o,x_{o'})} 4x_o + 2x_{o'}$ such that $p_o x_o + p_{o'} x_{o'} \leq 1$, where x_o and $x_{o'}$ are the probability shares of o and o' respectively. It concludes that the solution is $(x_o^\star, x_{o'}^\star) = (1,0)$ and thus purchases a unit of probability shares of o for player 1. The mechanism solves an analogous problem for player 3 and purchases a unit of probability shares of o' for player 3. Since player 2 has a budget of zero, the mechanism does not purchase any probability shares on her behalf. Since there is a unit supply of probability shares of each object, markets clear and the prices and shares constitute the Walrasian equilibrium. Since the equilibrium probability shares determine the assignment, with a unit share indicating a 100% chance of ownership, $M(\theta)$ assigns o to player 1, η to player 2, and o' to player 3.

It turns out that M is interim inefficient at any type profile where it is efficient for player 2 to receive an object. To illustrate, suppose nature draws B for player 1, A for player 2, and C for player 3, i.e., types are $\theta' = (B,A,C)$

[31] More recent treatments of Pseudo-markets and their derivative mechanisms can be found, for instance, in Budish (2011), Budish et al. (2013), and He et al. (2018).

instead of θ. Then, it is uniquely interim efficient for η to go to player 1, for o to go to player 2, and for o' to go to player 3. Yet, this assignment cannot be the outcome of $M(\theta')$ because budgets are pre-specified: simply, player 2 has a budget of zero, so if the mechanism is to purchase any probability shares of o on her behalf, it must be that $p_o = 0$. Yet, if $p_o = 0$, the mechanism would purchase infinite quantities of probability shares of o for all three players and markets would not clear. Hence, $M(\theta')$ is interim inefficient.

More generally, if is interim efficient for player i to receive an object at one type profile and to receive η under another type profile, then no Pseudo-market mechanism is interim efficient because efficiency requires i have a non-zero budget under the former profile and a zero budget under the latter profile, an impossibility. It is thus the fixed nature of budgets that drives the Walrasian equilibrium away from an interim efficient assignment, despite the fact the equilibrium consumes cardinal information. $\triangle$

In light of this example, one might consider developing an "alternative" Pseudo-market mechanism M by varying a Pseudo-market's budgets in response to players' reports and then allocating objects according to one of its equilibria. Miralles & Pycia (2014, Theorem 1) show that every Pareto efficient assignment corresponds to some Walrasian equilibrium of some Pseudo-market. Thus, for every profile of types θ, if one were to set the budgets in the right way, one could ensure that one of the resulting Walrasian equilibria was interim efficient and have $M(\theta)$ allocate objects according to this equilibrium. However, such a mechanism would be computationally difficult because there are no known, exact methods to compute either the necessary budgets or the equilibrium prices. We thus develop an alternative, interim efficient mechanism that is based on a linear programming problem.

5. THE SIMPLE MECHANISM

In the Illustration, we informally introduced the Simple Mechanism as a reaction to the interim inefficiency of other classic mechanisms and informally developed at its properties. Our goal, in this section, is to formally describe this mechanism and its core properties.

The Simple Mechanism.[32]

[32] This mechanism benefited greatly from discussions with Marek Pycia.

Suppose the players report $r = (r_1, \ldots, r_N) \in \Theta^N$. Then, the mechanism returns the mixed assignment $M_S(r)$, which is constructed as follows. First, for each player i, construct an estimate $\hat{v}_i$ of i's expected true values by setting

$$\hat{v}_i = (\hat{v}_{\eta i}, \hat{v}_{1i}, \ldots, \hat{v}_{Ki}) = \begin{cases} r_i & \text{if } r_i \in \mathscr{V} \\ \sum_{v \in I(r_i)} v \frac{f_v(v)}{\sum_{v \in I(r_i)} f_v(v)} & \text{if } r_i \in \mathscr{B}', \end{cases} \tag{4}$$

and represent $\hat{v}_i$ as a function $\hat{u}_i$ by setting $\hat{u}_i(o) = \hat{v}_{oi}$ for each $o \in \mathscr{O}'$. Second, compute the set

$$\sigma = \arg\max_{\phi \in \Phi} \sum_{i \in \mathscr{N}} \hat{u}_i(\phi(i)). \tag{5}$$

Third, set $M_S(r)$ to assign weight $1/|\sigma|$ to each pure assignment in σ.[33] ○

In words, the Simple Mechanism takes in players' reports, estimates their expected true values,[34] computes the set of pure assignments that maximize the sum of these estimated values, and then selects one of these assignments to implement with uniform probability.

The next two propositions formally establish the properties of the Simple Mechanism that we outlined in Observations 2 to 4: interim efficiency, symmetry, and ε-Bayesian incentive compatibility. All proofs are provided in the Appendix.

We say that a mechanism M is **symmetric** if, when players i and k make the same report $r_i = r_k$ and when they have the same type $\theta_i = \theta_k$, they have the same payoff to the (mixed) assignment $M(r)$, where $r = (r_1, \ldots, r_i, \ldots, r_k, \ldots, r_N) \in \Theta^N$.

Proposition 2. The Simple Mechanism is Interim Efficient and Symmetric. *The Simple Mechanism is interim efficient and symmetric.*

Interim efficiency implies that, under truthful reporting, the Simple Mechanism is interim Pareto efficient, ex-post Pareto efficient, and ordinally efficient

[33] Since Φ is finite, σ non-empty and finite.

[34] The Simple Mechanism requires knowledge of the structure of preferences to carry out its estimation; we examine the robustness of the mechanism with respect to this requirement in the Online Appendix. Also, a referee insightfully observed this requirement "shifts the [computational] burden" from players to the mechanism. From a mathematical perspective, one could shift it back by requiring players submit estimates of their expected true values to a mechanism, which assigns objects to maximize the sum of reported values. Provided this mechanism randomizes uniformly over the set of maximizing assignments, it has similar properties to the Simple Mechanism.

(when true values are strict); it thus is also individually rational.[35] Symmetry implies that the Simple Mechanism is "fair." The result's intuition mirrors the intuitions for Observations 2 and 3, and the proof amounts to a formalization of these intuitions.

Our incentive compatibility result relies on $\mathscr{V}$ having particular type of structure. Specifically, we say that $\mathscr{V}$ is a **product of finite grids** if $\mathscr{V} = \times_{o \in \mathscr{O}'} \mathscr{V}_o$, where $\mathscr{V}_o = \{v_1^o, v_2^o, \ldots, v_{L_o-1}^o, v_{L_o}^o\}$ for some $0 \leq v_1^o < v_2^o < \cdots < v_{L_o}^o \leq 1$ and $L_o \geq 1$ for each $o \in \mathscr{O}'$. For instance, if $\mathscr{V} = \{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}^{K+1}$ or if $\mathscr{V} = \{0\} \times \{0, \frac{1}{2}, 1\}^K$, then it is a product of finite grids. This mild restriction on $\mathscr{V}$ ensures that, with positive probability, we can find a player who has a higher value for one object and a lower value for another object. The restriction of $\mathscr{V}$ to the unit hyper-cube is without loss because we can always rescale the upper and lower bounds.

Proposition 3. The Simple Mechanism is ε-Bayesian Incentive Compatible. *Let $\mathscr{V}$ be a product of finite grids and let $\alpha > 0$. Then, for each $\varepsilon > 0$, there is a finite $\bar{N} > 0$ such that the Simple Mechanism is ε-Bayesian incentive compatible when there are at least $\bar{N}$ players.*

Since strategic manipulation is inherently costly, this proposition implies that the Simple Mechanism has good truth-telling properties when the number of players is large. The economic intuition is akin to the economic intuition behind Observation 3: as more players are added, "competition" for objects increases, which drives the probability a player can acquire any object, besides η, down to zero. This limits the opportunity for strategic manipulation and leads to ε-Bayesian incentive compatibility. Increasing competition itself is driven by the mechanism's randomization and the finite grid assumption, which ensures that, for each object, players with high values for the object enter and "drive up the demand" as N grows. The technical approach of the proof mirrors the argument used in Section 3, but differs at key points due to the generality of the environment.[36]

[35] It is evident that there are conditions under which the Simple Mechanism is "strictly" individually rational, i.e., all players do strictly better by participating in it than by remaining unassigned. For example, this is the case when (i) η is the least preferred option for all players and when (ii) there is at least one object that all players strictly prefer to η. It is also the case when (i') all players have the same most preferred object and (ii') value this object strictly more than η.

[36] The proof suggests a more general and abstract sufficient condition for ε-Bayesian incentive compatibility than finite grids, which we discuss in the Appendix.

A key assumption of the intuition (and of the proof) is that the set of objects $\mathscr{O}$ is held constant while N grows. To dispense with this assumption, suppose $\mathscr{O}$ grows via replication. A **linking function** is a $\psi : \mathbb{N}_+ \to \mathbb{N}_+$ that is monotone and has $\lim_{n\to\infty} \psi(n) = \infty$. We have in mind (i) that $\psi(N)$ gives the number of copies of each object and (ii) that players treat copies of objects as indistinguishable; we offer a formal treatment in the Appendix.

Proposition 4. Incentive Compatibility with Replication.
Let $\mathscr{V}$ be a product of finite grids and let $\alpha > 0$. For each $\varepsilon > 0$:
(i) When the linking function ψ is arbitrary, there may not exist a symmetric, interim efficient, and ε-Bayesian compatible mechanism.
(ii) There is a linking function $\bar{\psi}$ and an integer $\bar{N} > 0$ such that the Simple Mechanism is ε-Bayesian incentive compatible when there are at least $\bar{N}$ players.

The first part of the proposition is based on an example given in the Appendix. The example has $\psi(N) = \lfloor N/2 \rfloor$ copies of a single object and two equally likely types of players, A and B. The type As value the object at 1, the type Bs value it at $\frac{1}{2}$, and both value being unassigned at 0. Since the expected number of type As equals the number of objects, any interim efficient and symmetric mechanism (including the Simple Mechanism) essentially assign all copies of the object to the type As, leaving the type Bs unassigned. A "back of the envelope calculation" thus suggests that if a type B lies, then she obtains the object with probability $\frac{N/2}{N/2+1} = \frac{1}{1+2/N}$ (per symmetry), which converges to 1 as $N \to \infty$; whereas she gets nothing if she tells the truth. While actual mathematics are more involved, this simple computation lays bare the intuition and establishes that every type B gains from lying, even as N grows; economically speaking, it shows that replication can negate competition.

The second part of the proposition is a consequence of Proposition 3. To see this, suppose N starts growing from one and $\mathscr{O}$ starts off with one copy of each object. We wish to know at which value of N we can add a second copy of each object without compromising ε-Bayesian incentive compatibility. To find this value, assume for a moment that the number of copies does *not* change with N. Then, Proposition 3 tells us that there are thresholds N_1 and N_2 above which the Simple Mechanism is ε-Bayesian incentive compatible when $\mathscr{O}$ contains, respectively, one copy of each object or two copies of each object. Thus, when N is growing and $\mathscr{O}$ starts off with one copy of each object, (i) ε-Bayesian incentive compatible holds once N exceeds N_1 and

(ii) we can add a second copy of each object once N grows larger than N_2 without losing approximate incentive compatibility. Continuing this exercise defines thresholds $N_3, N_4, \ldots$ for three, four, or more copies and so implicitly defines the required linking function. Broadly, any linking function arrived at in this way has the number of copies increase sufficiently slowly with N so that competition increases (due to the use of Proposition 3).[37] (We provide a second example in the Appendix where we slow the replication in the prior example from $\lfloor N/2 \rfloor$ down to $\lfloor \ln(N)/2 \rfloor$ and establish that this is sufficient to ensure the Simple Mechanism is ε-Bayesian incentive compatible.)

A. APPENDIX: PROOFS AND ADDITIONAL EXAMPLES

This Appendix collects the proofs of our results, as well as additional examples, which are presented in the order that they are referenced in the main text.

Proof of Lemma 1. Obvious and omitted. □

Proof of Lemma 2. Obvious and omitted. □

Proof of Proposition 1. Let $\bar{N} = |\{\hat{o} \in \mathscr{O} | \check{v}_{\hat{o}} > \check{v}_\eta\}|$ and let $N > \bar{N}$. Our goal is to show that, M generates an interim inefficient assignment for some type profile.

To begin, suppose the type profile is $\theta = (\check{v}, \ldots, \check{v})$ and let ϕ be a pure assignment in the support of $M(\theta)$. There are two cases to consider: (i) either o or o' is unassigned by ϕ or (ii) both objects are assigned by ϕ.

For case (i), let $\tilde{o} \in \{o, o'\}$ denote an unassigned object. Since (i) there are $|\{\hat{o} \in \mathscr{O} | \check{v}_{\hat{o}} > \check{v}_\eta\}|$ objects that all players value more than being unassigned and (ii) since $N > |\{\hat{o} \in \mathscr{O} | \check{v}_{\hat{o}} > \check{v}_\eta\}|$, at least one player is either assigned η or is assigned some other object o'' that she values no more than η. Let i denote such a player. Construct a new assignment ϕ' by giving $\tilde{o}$ to player i, while leaving everyone else's assignment alone – i.e., set $\phi'(i) = \tilde{o}$ and $\phi'(k) = \phi(k)$ for all players $k \neq i$. Observe that $\sum_{l \in \mathscr{N}} u_l(\phi' | \theta_l) - \sum_{l \in \mathscr{N}} u_l(\phi | \theta_l) = u_i(\phi' | \theta_i) - u_i(\phi | \theta_i) = \check{v}_{\tilde{o}} - u_i(\phi | \theta_i) > 0$ since $\tilde{o}$ is strictly preferred to η. Thus, ϕ is interim inefficient. Since a mixed assignment is interim efficient if and

[37] It is also possible to obtain approximate incentive compatibility by having the number of copies increase sufficiently quickly with N such that scarcity "evaporates" and everyone gets their most preferred object – e.g., set $\psi(N) \geq N$ for all N. However, such linking functions are economically uninteresting.

only if each pure assignment in its support is interim efficient, it follows that $M(\theta)$ is also interim inefficient.

For case (ii), let i be the player who gets o and let k be the player who gets o'. Consider a change in i's type from $\check{v}$ to $\check{v}'$. Formally, consider a new type profile $\theta' = (\theta'_1, \ldots, \theta'_{i-1}, \theta'_i, \theta'_{i+1}, \ldots, \theta'_N)$, where $\theta'_i = \check{v}'$ and $\theta'_l = \check{v}$ for all players $l \neq i$; this type profile occurs with positive probability per Assumption 1 and the fact $\alpha > 0$.

Since (i) M disregards cardinal information and (ii) $c(\theta') = c(\theta)$, we have $M(\theta') = M(c(\theta')) = M(c(\theta)) = M(\theta)$. Thus, ϕ is in the support of $M(\theta')$. Construct a new assignment ϕ' from ϕ by swapping i and k's assignments, while leave all other players' assignments alone – i.e., set $\phi'(i) = o''$, $\phi'(k) = o'$, and $\phi'(l) = \phi(l)$ for all players $l \neq i, k$. We have

$$\sum_{l \in \mathcal{N}} u_l(\phi'|\theta'_l) - \sum_{l \in \mathcal{N}} u_l(\phi|\theta'_l) = u_i(\phi'|\theta'_i) - u_i(\phi|\theta'_i) + u_k(\phi'|\theta'_k) - u_k(\phi|\theta'_k)$$
$$= \check{v}'_{o'} - \check{v}'_{o''} - (\check{v}_{o'} - \check{v}_{o''}) > 0,$$

where the strict inequality follows from Assumption 1. That is, ϕ is interim inefficient, implying $M(\theta')$ is interim inefficient. □

Example 1 Details. The Nash Boston Symmetric Equilibrium.

In the Boston Mechanism's reporting game, players submit (strict) ordinal reports.[38] The mechanism then allocates objects by sequentially considering all players' first choices (in the first round), second choices (in the second round), and so on; it stops once each player is assigned an object or η. In the k-th round, the mechanism first "pins" the name of each player (who was not assigned to an object or η in a previous round) to her k-th choice from $\mathscr{O}'$, provided this choice is available (i.e., was not given to another player in a previous round); η is always available. If an object has one player's name pinned to it, then it is assigned to that player. However, if an object has two or more players' names pinned to it, then a (uniform) random lottery is held and the object is assigned to the winner. The item η is assigned to all players whose names are pinned to it. Players who lose their lotteries, players who have k-th choices that are unavailable, and objects that had no names pinned to them all move forwards into the next round.

As is common in the literature on the Boston mechanism (e.g., Abdulkadiroglu et al., 2011), we look for a symmetric Nash equilibrium. Observe that

[38] Abdulkadiroglu & Sonmez (2003) provide a detailed description of the Boston mechanism, which we have adapted for our single-unit-demand environment.

if a player's k-th choice is η and they enter the k-th round, then the mechanism assigns them η for certain. Thus, it is dominated for a type A to report o below η or to report o' above η – so, her dominant strategy is $o \succ \eta \succ o'$ – and it is dominated for a type B to report o or o' below η – so, she has two possible strategies: $s_1 := o \succ o' \succ \eta$ or $s_2 := o' \succ o \succ \eta$.[39]

If B plays s_1, then with probability $\frac{1}{2}$ she faces a type A. Since both players list o first, the mechanism conducts a fair lottery for it in round one; if the type B wins she gets it and if she loses then she gets o' in the next round. Hence, her payoff to facing a type A is $\frac{9}{2} + \frac{8}{2}$. A similar argument gives that her payoff to facing a type B is $\frac{9}{2} + \frac{8}{2}$ because of symmetry. Thus, her payoff to s_1 is $\frac{1}{2}(\frac{9}{2} + \frac{8}{2}) + \frac{1}{2}(\frac{9}{2} + \frac{8}{2})$. If B plays s_2, then her payoff is $\frac{1}{2}8 + \frac{1}{2}(\frac{9}{2} + \frac{8}{2})$ via analogous logic. Thus, a type B's best response is s_1.

It follows that the symmetric Nash equilibrium is $o \succ \eta \succ o'$ if a type A and $o \succ o' \succ \eta$ if a type B. Computation then gives the mixed assignment for each type-profile reported in the main text. △

Proof of Proposition 2. Suppose the players types are $\theta = (\theta_1, \ldots, \theta_N)$, so their expected true values are $\{v_i^\dagger\}_{i \in \mathcal{N}}$ per equation (2) and their payoffs are $\{u_i\}_{i \in \mathcal{N}}$. We first wish to show that $M_S(\theta)$ is interim efficient. When the Simple Mechanism receives report θ, it computes estimates of the players' expected true values $\{\hat{v}_i\}_{i \in \mathcal{N}}$ according to equation (4) and payoffs $\{\hat{u}_i\}_{i \in \mathcal{N}}$. Since equation (4) is identical to equation (2), we have $\hat{v}_i = v_i^\dagger$ for each player i. Thus, when the Simple Mechanism assigns objects to maximize its estimate of social surplus (i.e., to solve $\max_{\phi \in \Phi} \sum_{i \in \mathcal{N}} \hat{u}_i(\phi(i))$), it is actually maximizing social surplus (i.e., solving $\max_{\phi \in \Phi} \sum_{i \in \mathcal{N}} u_i(\phi(i)|\theta_i)$). Thus, the mixed assignment it produces is interim efficient.

We next wish to show that the Simple Mechanism is symmetric. Let i and k be two players who make the same report r and have the same type θ; subscripts are suppressed for simplicity. Upon receiving report $r = (r_1, \ldots, r_{i-1}, r, r_{i+1}, \ldots, r_{k-1}, r, r_{k+1}, \ldots, r_N)$, the Simple Mechanism generates estimates of players' expected true values $\{\hat{v}_l\}_{l \in \mathcal{N}}$ according to equation (4) and estimates of their payoffs $\{\hat{u}_l\}_{l \in \mathcal{N}}$. It also computes $\sigma = \arg\max_{\phi \in \Phi} \sum_{i \in \mathcal{N}} \hat{u}_l(\phi(l))$.

Partition σ as follows $\sigma_+ = \{\phi \in \sigma | u_i(\phi|\theta) > u_k(\phi|\theta)\}$, $\sigma_= = \{\phi \in \sigma | u_i(\phi|\theta) = u_k(\phi|\theta)\}$, and $\sigma_- = \{\phi \in \sigma | u_i(\phi|\theta) < u_k(\phi|\theta)\}$. Focus on the relationship between σ_+ and σ_-. For each $\phi \in \sigma_+$, we may construct an

[39] We write $x \succ y$ to mean x is strictly preferred to y.

element $s(\phi) \in \sigma_-$ by having i and k swap their assignments, while leaving all others players' assignments unaltered, i.e., by setting $s(\phi)(i) = \phi(k)$, $s(\phi)(k) = \phi(i)$, and $s(\phi)(l) = \phi(l)$ for every other player l. This swap: (i) does not change the Simple Mechanism's estimate of the sum of players' payoffs since the mechanism estimates i and k have the same payoff (since they make the same report) and (ii) defines a bijection between σ_+ and σ_-.

Turning to i and k payoffs to the Simple Mechanism, we have $u_i(M_S(r)|\theta) =$

$$\begin{aligned}
& \sum_{\phi\in\sigma_+} u_i(\phi|\theta)\frac{1}{|\sigma|} + \sum_{\phi\in\sigma_=} u_i(\phi|\theta)\frac{1}{|\sigma|} + \sum_{\phi\in\sigma_-} u_i(\phi|\theta)\frac{1}{|\sigma|} \\
= & \sum_{\phi\in\sigma_+} u_k(s(\phi)|\theta)\frac{1}{|\sigma|} + \sum_{\phi\in\sigma_=} u_k(\phi|\theta)\frac{1}{|\sigma|} + \sum_{\phi\in\sigma_-} u_k(s(\phi)|\theta)\frac{1}{|\sigma|} \\
= & \sum_{\phi\in\sigma_-} u_k(\phi|\theta)\frac{1}{|\sigma|} + \sum_{\phi\in\sigma_=} u_k(\phi|\theta)\frac{1}{|\sigma|} + \sum_{\phi\in\sigma_+} u_k(\phi|\theta)\frac{1}{|\sigma|} = u_k(M_S(r)|\theta).
\end{aligned}$$

The second line follows from the fact that i and k swap assignments and therefore payoffs. The third line follows from reindexing: since s is bijective, we may replace each $\phi \in \sigma_+$ ($\phi \in \sigma_-$) with its corresponding element from σ_- (σ_+), implying $\sum_{\phi\in\sigma_+} u_k(s(\phi)|\theta) = \sum_{\phi\in\sigma_-} u_k(\phi|\theta)$ and $\sum_{\phi\in\sigma_-} u_k(s(\phi)|\theta) = \sum_{\phi'\in\sigma_+} u_k(\phi|\theta)$. □

We prove Proposition 3 in two steps. First, we establish an upper bound on any player's payoff to sending a report to the Simple Mechanism, when the other players tell the truth, as a function of her type and the number of players. Then, we show that this bound goes to her value of η as N gets large – Lemma A1. Second, we leverage this bound, the fact that a strategic liar never does worse than being unassigned, and a squeeze argument (similar to the one used to establish Observation 4) to show that each player's gain from strategic lying is small when N is large.

Lemma A1. A Uniform Upper Bound on $U_i^{M_S}(r_i|\theta_i)$.
Let $\mathscr{V}$ be the product of finite grids and let $\alpha > 0$. Then, there is a function $g : \Theta \times \mathbb{N} \to \mathbb{R}$, such that, for each player i and each $\theta_i \in \Theta$:
(i) The function g bounds i's payoff to reporting each $r_i \in \Theta$ from above, i.e., $U_i^{M_S}(r_i|\theta_i) \leq g(\theta_i, N)$.
(ii) The function g converges to i's value of being unassigned as $N \to \infty$, i.e., $\lim_{N\to\infty} g(\theta_i, N) = u_i(\eta|\theta_i)$.

Proof. Consider player i. We proceed in four steps. First, we partition the type space of the other players in the right way. Second, on each of these partitions, we give upper bounds for i's payoff given the behavior of the Simple Mechanism. Third, we average these upper bounds and write down a closed form expression of g. Fourth, we let $N \to \infty$ and show that g takes the claimed limits.

We begin by partitioning the type space Θ^{N-1} of the other players. We say a player is a **type** E if her type is $(v_1^\eta, v_{L_1}^1, v_{L_2}^2, \ldots, v_{L_K}^K)$, i.e., if she (i) knows her true values, (ii) values being unassigned at the lowest possible value, and (iii) values every object at the highest possible value. (There are such players because $\mathscr{V}$ is the product of finite grids, $\alpha > 0$, and $f_v > 0$.) Let E_l denote the event that there are l type E players besides i, i.e., $E_l = \{\theta_{-i} \in \Theta^{N-1} | \theta_{-i}$ specifies l players have type $E\}$. The probability nature draws type E for a given player is $\alpha f_v(\bar{v}) > 0$, where $\bar{v} = (v_1^\eta, v_{L_1}^1, v_{L_2}^2, \ldots, v_{L_K}^K)$. Thus, the probability of E_l is $\Pr(E_l) = \sum_{\theta_{-i} \in E_l} \Pr(\theta_{-i}) = \binom{N-1}{l}(\alpha f_v(\bar{v}))^l (1 - \alpha f_v(\bar{v}))^{N-1-l}$ due to independence. By the Binomial Theorem, $\sum_{l=0}^{N-1} \Pr(E_l) = 1$.

Next, we give upper bounds on $u_i(M_S(r_i, \theta_{-i}) | \theta_i)$ on each set E_l for each $r_i \in \Theta$ and each $\theta_i \in \Theta$. For l between 0 and K, it suffices to let 1 be this bound by the finite grid assumption. For $l > K$, we will establish that a bound is $u_i(\eta | \theta_i) + \frac{K}{l-K}$. It follows that

$$
\begin{aligned}
U_i^M(r_i|\theta_i) &= \sum_{\theta_{-i} \in \Theta^{N-1}} u_i(M_S(r_i, \theta_{-i})|\theta_i) \Pr(\theta_{-i}) \\
&= \sum_{l=0}^{K} \sum_{\theta_{-i} \in E_l} u_i(M_S(r_i, \theta_{-i})|\theta_i) \Pr(\theta_{-i}) \\
&\quad + \sum_{l=K+1}^{N-1} \sum_{\theta_{-i} \in E_l} u_i(M_S(r_i, \theta_{-i})|\theta_i) \Pr(\theta_{-i}) \\
&\leq \sum_{l=0}^{K} \sum_{\theta_{-i} \in E_l} 1 \Pr(\theta_{-i}) + \sum_{l=K+1}^{N-1} \sum_{\theta_{-i} \in E_l} (u_i(\eta|\theta_i) + \frac{K}{l-K}) \Pr(\theta_{-i}) \\
&= \sum_{l=0}^{K} \Pr(E_l) + \sum_{l=K+1}^{N-1} (u_i(\eta|\theta_i) + \frac{K}{l-K}) \Pr(E_l)
\end{aligned}
$$

for each $r_i \in \Theta$ and each $\theta_i \in \Theta$. We take the last line to be our upper bound g,

i.e., we set

$$g(\theta_i, N) = \sum_{l=0}^{K} \Pr(E_l) + \sum_{l=K+1}^{N-1} (u_i(\eta|\theta_i) + \frac{K}{l-K}) \Pr(E_l)$$

for each $\theta_i \in \Theta$. (The bound applies to every player, including player i, because players with the same type have the same payoff.)

We need to show that g takes the claimed limits. We begin by noting $\lim_{N\to\infty} \Pr(E_l) = 0$ for all $l < \infty$.[40] Thus, $\sum_{l=0}^{K} \Pr(E_l) \to 0$ as $N \to \infty$. To show that $\sum_{l=K+1}^{N-1} (u_i(\eta|\theta_i) + \frac{K}{l-K}) \Pr(E_l) \to u_i(\eta|\theta_i)$ for each $\theta_i \in \Theta$, we use a squeeze argument. Let $\delta > 0$, then there is a fixed N' such that (i) $\frac{K}{l-K} > \delta$ for all $l < N'$ and (ii) $\frac{K}{l-K} \leq \delta$ for all $l \geq N'$. Thus, for large N, we have

$$\begin{aligned}
&\sum_{l=K+1}^{N-1} (u_i(\eta|\theta_i) + \frac{K}{l-K}) \Pr(E_l) \\
&\leq u_i(\eta|\theta_i) + \sum_{l=K+1}^{N'-1} \frac{K}{l-K} \Pr(E_l) + \delta \sum_{l=N'}^{N-1} \Pr(E_l) \\
&= \underbrace{u_i(\eta|\theta_i) + \sum_{l=K+1}^{N'-1} \frac{K}{l-K} \Pr(E_l) + \delta(1 - \sum_{l=0}^{N'-1} \Pr(E_l))}_{h(N)}.
\end{aligned}$$

Since $\lim_{N\to\infty} \Pr(E_l) = 0$, the finite sums in $h(N)$ go to zero as $N \to \infty$, implying $h(N) \to u_i(\eta|\theta_i) + \delta$. Thus, we have $\lim_{N\to\infty} g(\theta_i, N) \leq u_i(\eta|\theta_i) + \delta$ for

[40] This follows from Sterling's Approximation and l'Hopital's rule. Let $p = \alpha f_v(\bar{v})$. For each finite l, Sterling's Approximation (see Feller, 1968, p.52) lets us write

$$\lim_{N\to\infty} \Pr(E_l) = \lim_{N\to\infty} (N-1)^l (1-p)^{N-1} (1 + \frac{l}{N-1-l})^{N-1-l} \sqrt{\frac{N-1}{N-1-l}} \frac{p^l}{l!\, e^l (1-p)^l}.$$

Since $\lim_{N\to\infty} (1 + \frac{l}{N-1-l})^{N-l-1} = e^l$, $\lim_{N\to\infty} \sqrt{\frac{N-1}{N-1-l}} = 1$, and $\frac{p^l}{l!\, e^l (1-p)^l}$ is constant, the limit of $\Pr(E_l)$ is zero if and only if $\lim_{N\to\infty} (N-1)^l (1-p)^{N-1} = 0$. Let $\beta = 1/(1-p)$, so $(N-1)^l (1-p)^{N-1} = (N-1)^l / \beta^{N-1}$. Since $\beta > 1$ as $1 - p \in (0,1)$, we need to apply l'Hopital's rule. Doing so l times to gives

$$\lim_{N\to\infty} \frac{(N-1)^l}{\beta^N} = \lim_{N\to\infty} \frac{l!\,(N-1)^0}{\frac{(N-1)!}{(N-l-2)!} \beta^{N-1-l}}.$$

It is evident that the right-hand-side has limit zero, implying $\lim_{N\to\infty} (N-1)^l (1-p)^{N-1} = 0$.

all $\delta > 0$. Since $g(\theta_i, N) \geq u_i(\eta|\theta_i)$, it follows that $\lim_{N\to\infty} g(\theta_i, N) = u_i(\eta|\theta_i)$ for each $\theta_i \in \Theta$.

It only remains to show that, when $l > K$,

$$u_i(M_S(r_i, \theta_{-i})|\theta_i) \leq u_i(\eta|\theta_i) + \frac{K}{l-K}, \tag{6}$$

for each $r_i \in \Theta$ and each $\theta_{-i} \in E_l$. To do this, we need four facts.

Fact 0. For each $r_i \in \Theta$, for each $\theta_{-i} \in E_l$, and for each ϕ in the support of $M_S(r_i, \theta_{-i})$, there are at least $l - K$ unassigned type E players.

Proof. Simply there are K objects and $l > K$ type E players. Thus, at least $l - K$ of the type E players must be unassigned. △

Fact 1. For each $r_i \in \Theta$ and each $\theta_{-i} \in E_l$, the Simple Mechanism assigns i an object $o \in \mathscr{O}$ with strictly positive probability only if it estimates that her value for said object is $v^o_{L_o}$. Thus, the Simple Mechanism assigns i to η if it estimates that her value for each object $o \in \mathscr{O}$ is strictly less than the maximum value for that object $v^o_{L_o}$.

Proof. Upon receiving report $r = (r_i, \theta_{-i})$, the Simple Mechanism generates estimates of players' expected true values $\{\hat{v}_l\}_{l\in\mathscr{N}}$ according to equation (4) and estimates of their payoffs $\{\hat{u}_l\}_{l\in\mathscr{N}}$. It also computes $\sigma = \arg\max_{\phi\in\Phi} \sum_{l\in\mathscr{N}} \hat{u}_l(\phi(l))$. Observe (i) that $\hat{u}_l(o) = u_l(o|\theta_l)$ for each player $l \neq i$ and each $o \in \mathscr{O}'$ per truthful reporting and (ii) that $v^o_1 \leq \hat{u}_l(o) \leq v^o_{L_o}$ for each player l and each $o \in \mathscr{O}'$.[41]

We argue the fact by contradiction. Suppose i is assigned an object $o \in \mathscr{O}$ with strictly positive probability and the mechanism estimates that $\hat{u}_i(o) < v^o_{L_o}$. Then, there is a pure assignment ϕ in the support of $M_S(r_i, \theta_{-i})$ where i gets o. Since there is at least one unassigned type E under ϕ (by Fact 0), call her k, we can construct a new assignment ϕ' from ϕ by assigning o to k, by assigning η to i, and by leaving all other players' assignments alone – formally, $\phi'(k) = o$, $\phi'(i) = \eta$, and $\phi'(l) = \phi(l)$ for all $l \neq i, k$. We will show that $\sum_{l\in\mathscr{N}} \hat{u}_l(\phi'(l)) -$

[41] Since equation (4) either returns an element of $\mathscr{V}$ or averages over elements of $\mathscr{V}$ and since every element of $\mathscr{V}$ is bounded below by $(v^\eta_1, \ldots, v^K_1)$ and above by $(v^\eta_{L_\eta}, \ldots, v^K_{L_K})$ as it is the product of finite grids, we necessarily have $(v^\eta_1, \ldots, v^K_1) \leq \hat{v}_l \leq (v^\eta_{L_\eta}, \ldots, v^K_{L_K})$ for each player l.

$\sum_{l\in\mathcal{N}} \hat{u}_l(\phi(l)) > 0$. Thus, ϕ does not maximize the Simple Mechanism's estimate of social welfare and therefore cannot be in the support of $M_S(r_i, \theta_{-i})$, a contradiction.

It remains to show that $\sum_{l\in\mathcal{N}} \hat{u}_l(\phi'(l)) - \sum_{l\in\mathcal{N}} \hat{u}_l(\phi(l)) > 0$. Since (i) $\hat{u}_i(o) < v^o_{L_o}$ and (ii) $\hat{u}_i(\eta) \geq v^\eta_1$, we have $v^o_{L_o} - \hat{u}_i(o) + \hat{u}_i(\eta) - v^\eta_1 > 0$. Since k is a truthful type E, we have $\hat{u}_k(o) - \hat{u}_k(\eta) = v^o_{L_o} - v^\eta_1$. Hence,

$$\begin{aligned}\sum_{l\in\mathcal{N}} \hat{u}_l(\phi'(l)) - \sum_{l\in\mathcal{N}} \hat{u}_l(\phi(l)) &= \hat{u}_k(o) - \hat{u}_k(\eta) + \hat{u}_i(\eta) - \hat{u}_i(o) \\ &= v^o_{L_o} - \hat{u}_i(o) + \hat{u}_i(\eta) - v^\eta_1 > 0. \triangle\end{aligned}$$

Fact 2. For each $r_i \in \Theta$ and each $\theta_{-i} \in E_l$, the Simple Mechanism assigns η to i if it estimates that her value of η is strictly greater than v^η_1.

Proof. Upon receiving report $r = (r_i, \theta_{-i})$, the Simple Mechanism generates estimates of players' expected true values $\{\hat{v}_l\}_{l\in\mathcal{N}}$ according to equation (4) and estimates of their payoffs $\{\hat{u}_l\}_{l\in\mathcal{N}}$. It also computes $\sigma = \arg\max_{\phi\in\Phi} \sum_{l\in\mathcal{N}} \hat{u}_l(\phi(l))$. Recall (i) that $\hat{u}_l(o) = u_l(o|\theta_l)$ for each player $l \neq i$ and each $o \in \mathcal{O}'$ since these players report truthfully and (ii) that $v^o_1 \leq \hat{u}_l(o) \leq v^o_{L_o}$ for each player l and each $o \in \mathcal{O}'$.

We argue this fact by contradiction using the same approach as in the proof of Fact 1. Suppose i is assigned an object $o \in \mathcal{O}$ with strictly positive and the mechanism estimates that $\hat{u}_i(\eta) > v^\eta_1$. Then, there is a pure assignment ϕ in the support of $M_S(r, \theta_{-i})$ where i gets o. Since there is at least one unassigned type E under ϕ, call her k, we can construct a new assignment ϕ' from ϕ by assigning o to k, by assigning η to i, and by leaving all other players' assignments alone – formally, $\phi'(k) = o$, $\phi'(i) = \eta$, and $\phi'(l) = \phi(l)$ for all $l \neq i, k$. We have

$$\begin{aligned}\sum_{l\in\mathcal{N}} \hat{u}_l(\phi'(l)) - \sum_{l\in\mathcal{N}} \hat{u}_l(\phi(l)) &= \hat{u}_k(o) - \hat{u}_k(\eta) + \hat{u}_i(\eta) - \hat{u}_i(o) \\ &= v^o_{L_o} - v^\eta_1 + \hat{u}_i(\eta) - \hat{u}_i(o),\end{aligned}$$

because $\hat{u}_k(o) - \hat{u}_k(\eta) = v^o_{L_o} - v^\eta_1$. Since $\hat{u}_i(o) \leq v^o_{L_o}$ and $\hat{u}_i(\eta) > v^\eta_1$, we have $v^o_{L_o} - \hat{u}_i(o) + \hat{u}_i(\eta) - v^\eta_1 > 0$. Thus, ϕ cannot be an element of $M_S(r_i, \theta_{-i})$, a contradiction. $\triangle$

Fact 3. For each $r_i \in \Theta$ and each $\theta_{-i} \in E_l$, if the Simple Mechanism assigns i an object $o \in \mathscr{O}$ with strictly positive probability, then the probability i gets o is bounded above by $\frac{1}{l-K}$. (Notice that this bound does *not* apply to η.)

Proof. Upon receiving report $r = (r_i, \theta_{-i})$, the Simple Mechanism generates estimates of players' expected true values $\{\hat{v}_l\}_{l \in \mathscr{N}}$ according to equation (4) and estimates of their payoffs $\{\hat{u}_l\}_{l \in \mathscr{N}}$. It also computes $\sigma = \arg\max_{\phi \in \Phi} \sum_{l \in \mathscr{N}} \hat{u}_l(\phi(l))$. Recall (i) that $\hat{u}_l(o) = u_l(o|\theta_l)$ for each player $l \neq i$ and each $o \in \mathscr{O}'$ and (ii) that $v_1^o \leq \hat{u}_l(o) \leq v_{L_o}^o$ for each player l and each $o \in \mathscr{O}'$.

Let $p_i(o)$ be the probability that i is assigned object $o \in \mathscr{O}$ in $M_S(r, \theta_{-i})$. Since the Simple Mechanism randomizes uniformly over σ, we have $p_i(o) = \sum_{\phi \in C_i(o)} \frac{1}{|\sigma|} = \frac{|C_i(o)|}{|\sigma|}$ where $C_i(o) = \{\phi \in \sigma | \phi(i) = o\}$ is the sub-set of σ where i gets o. (Since $p_i(o) > 0$ by hypothesis, $C_i(o)$ is non-empty.) Our goal is to show that $p_i(o) \leq \frac{1}{l-K}$, i.e., $|C_i(o)|(l-K) \leq |\sigma|$. We do this by showing that for each $\phi \in C_i(o)$, there are $l-K$ "unique" assignments in $\sigma \backslash C_i(o)$.

Pick a $\phi \in C_i(o)$ and let $S(\phi)$ be the set of unassigned type E players at ϕ. For each $k \in S(\phi)$ we can create an assignment ϕ' by swapping i and k's assignments and leaving all other players' assignments alone, i.e., by setting $\phi'(k) = o$, $\phi'(i) = \eta$, and $\phi'(l) = \phi(l)$ for all $l \neq i, k$. Since there are at least $l-K$ players in $S(\phi)$ per Fact 0, we can create at least $l-K$ different assignments in this way. Let $P(\phi)$ be the set of assignments created in this fashion.

First, observe that $P(\phi) \subset \sigma \backslash C_i(o)$. It is trivial that $P(\phi) \subset \Phi \backslash C_i(o)$, since i is assigned η in every element of $P(\phi)$. So, we only need to show that $P(\phi) \subset \sigma$. This follows from Facts 1 and 2: since i receives object o with strictly positive probability in $M_S(r, \theta_{-i})$, it must be that $\hat{u}_i(\eta) = v_1^\eta$ and $\hat{u}_i(o) = v_{L_o}^o$. Thus, $\sum_{l \in \mathscr{N}} \hat{u}_l(\phi'(l)) - \sum_{l \in \mathscr{N}} \hat{u}_l(\phi(l)) = \hat{u}_{\phi'(o)}(o) - \hat{u}_{\phi'(o)}(\eta) + \hat{u}_i(\eta) - \hat{u}_i(o) = 0$ for any $\phi' \in P(\phi)$ since $\hat{u}_{\phi'(o)}(o) = v_{L_o}^o$ and $\hat{u}_{\phi'(o)}(\eta) = v_1^\eta$ as $\phi'(o)$ is truthful type E. It follows that $P(\phi) \subset \sigma$.

Second, observe that for two assignments ϕ and ϕ' in $C_i(o)$, we have that $P(\phi) \cap P(\phi') = \emptyset$. We argue this by contradiction. Suppose there is a $\tilde{\phi} \in P(\phi) \cap P(\phi')$. Since ϕ and ϕ' are different assignments, there is a player k for whom $\phi(k) \neq \phi'(k)$. (Notice $k \neq i$ as i has object o in both ϕ

and ϕ'.) There are three cases: $\tilde{\phi}(k) = o'$ for some $o' \in \mathscr{O}\backslash\{o\}$, $\tilde{\phi}(k) = \eta$, and $\tilde{\phi}(k) = o$. If $\tilde{\phi}(k) = o'$, then it must be that $\phi(k) = \phi'(k) = o'$ – since we do not chance the assignments of players who get objects other than o in the creation of $\tilde{\phi}$, the only way k could end up with o' is if she had it to begin with in both ϕ and ϕ'. If $\tilde{\phi}(k) = \eta$, then it must be that $\phi(k) = \phi'(k) = \eta$ since we never assign η to a player (except i) in the creation of $\tilde{\phi}$, unless said player has η to begin with. If $\tilde{\phi}(k) = o$, then it must be that $\phi(k) = \phi'(k) = \eta$ since we only give o to an unassigned type E player in the creation of $\tilde{\phi}$. In all three cases, $\phi(k) = \phi'(k)$, a contradiction.

Let $Q = \cup_{\phi \in C_i(o)} P(\phi)$. Since the $\{P(\phi)\}_{\phi \in C_i(o)}$ sets are disjoint by the second observation, $|Q| = \sum_{\phi \in C_i(o)} |P(\phi)|$. Since $|P(\phi)| \geq (l-K)$, we have $|Q| \geq (l-K)|C_i(o)|$. Since $Q \subset \sigma\backslash C_i(o)$ by the first observation, σ has at least $(l-K)|C_i(o)|$ elements. $\triangle$

This concludes the list of facts and the proofs thereof.

We now establish equation (6). Let $l > K$, let $\theta_{-i} \in E_l$, and, for each $r_i \in \Theta$, let $\hat{u}_i^{r_i}$ be the Simple Mechanism's estimate of i's payoff function upon receiving the report (r_i, θ_{-i}), see equation (4). We partition the set of all possible reports Θ into three parts Θ_1, Θ_2, and Θ_3 as follows: $\Theta_1 = \{r_i \in \Theta | \hat{u}_i^{r_i}(\eta) > v_1^\eta\}$, $\Theta_2 = \{r_i \in \Theta | \hat{u}_i^{r_i}(\eta) = v_1^\eta\} \cap \{r_i \in \Theta | \hat{u}_i^{r_i}(o) < v_{L_o}^o$ for all $o \in \mathscr{O}\}$, and $\Theta_3 = \{r_i \in \Theta | \hat{u}_i^{r_i}(\eta) = v_1^\eta\} \cap \{r_i \in \Theta | \hat{u}_i^{r_i}(o) = v_{L_o}^o$ for some $o \in \mathscr{O}\}$. In other words, Θ_1 is the set of reports that cause the mechanism to estimate that i's value to η is strictly greater than v_1^η, Θ_2 is the set of reports that cause the mechanism to estimate that i's value to η is v_1^η and her value to every object o is less than $v_{L_0}^o$, and Θ_3 is the set of reports cause the mechanism to estimate that i's value to η is v_1^η and i's value to some objects is maximal.

If $r_i \in \Theta_1 \cup \Theta_2$, then Facts 1 and 2 tells us the Simple Mechanism assigns η to i, so $u_i(M_S(r_i, \theta_{-i})|\theta_i) = u_i(\eta|\theta_i)$. If $r_i \in \Theta_3$, then Fact 3 tells us that the probability i gets each object in $\mathscr{O}$ is no more than $\frac{1}{l-K}$. Hence,

$$\begin{aligned} u_i(M_S(r_i, \theta_{-i})|\theta_i) &= u_i(\eta|\theta_i)p_i(\eta) + \sum_{o \in \mathscr{O}} u_i(o|\theta_i)p_i(o) \\ &\leq u_i(\eta|\theta_i) + \sum_{o \in \mathscr{O}} u_i(o|\theta_i)\frac{1}{l-K} \leq u_i(\eta|\theta_i) + \frac{K}{l+1} \end{aligned}$$

for any $r_i \in \Theta_3$, where $p_i(o)$ is the probability i gets o in $M_S(r_i, \theta_{-i})$ and the last inequality is due to the finite grid assumption. Thus, across all reports

in $r_i \in \Theta$, we have $u_i(M_S(r_i, \theta_{-i})|\theta_i) \leq u_i(\eta|\theta_i) + \frac{K}{l-K}$. That is, equation (6) holds. □

Proof of Proposition 3. Consider player i whose type is $\theta_i \in \Theta$. We begin by establishing that $u_i(\eta|\theta_i)$ is a lower bound on i's earnings when she strategically lies or when she reports truthfully. Observe that $u_i(M_S(\theta_i, \theta_{-i})|\theta_i) \geq u_i(\eta|\theta_i)$ for all $\theta_{-i} \in \Theta^{N-1}$ by interim efficiency and individual rationality. It follows that $U_i^{M_S}(\theta_i|\theta_i) =$

$$\sum_{\theta_{-i}\in\Theta^{N-1}} u_i(M_S(\theta_i, \theta_{-i})|\theta_i)\Pr(\theta_{-i}) \geq \sum_{\theta_{-i}\in\Theta^{N-1}} u_i(\eta|\theta_i)\Pr(\theta_{-i}) = u_i(\eta|\theta_i).$$

Hence, $\max_{r_i\in\Theta} U_i^{M_S}(r_i|\theta_i) \geq U_i^{M_S}(\theta_i|\theta_i) \geq u_i(\eta|\theta_i)$ by optimality.

Now, let $\varepsilon > 0$. Then, by the second part of Lemma A1, there is a finite $N_{\theta_i} > 0$ such that $g(\theta_i, N) \leq u_i(\eta|\theta_i) + \varepsilon$ for all $N \geq N_{\theta_i}$. The first part of Lemma A1 and the prior paragraph then imply $u_i(\eta|\theta_i) \leq U_i^{M_S}(\theta_i|\theta_i) \leq u_i(\eta|\theta_i) + \varepsilon$ and $u_i(\eta|\theta_i) \leq \max_{r_i\in\Theta} U_i^{M_S}(r_i|\theta_i) \leq u_i(\eta|\theta_i) + \varepsilon$ for all $N \geq N_{\theta_i}$, i.e., i never gain more than ε by lying strategically when there are at least N_{θ_i} players. Thus, no player gains more than ε by strategically lying when there are at least $\bar{N} = \max_{\theta\in\Theta} N_\theta$ players (since players with the same type have the same payoff.) □

Remark 1. It is clear from the Proofs of Lemma A1 and Proposition 3 that we can relax the assumption that $\mathscr{V}$ is a product of finite grids to the assumption that, for each $o \in \mathscr{O}$, (i) there is a type of player who has the maximum value for o (among all players) and the minimum value for η and (ii) this player occurs with non-zero probability. We omit this abstract generalization because it adds little economic insight.

To extend the baseline model (of Section 2) to allow for object replication, we define the **replication extension** using the baseline model's primitives – $\mathscr{O}$, $\mathscr{V}$, f_v, α, and N – as well as the linking function $\psi(N)$. This extension makes two key modifications to the baseline model. First, the set of objects is $\cup_{l=1}^{\psi(N)}\mathscr{O}$ and a pure assignment ϕ is a map from $\mathscr{N}$ into $\{\eta\} \cup_{l=1}^{\psi(N)} \mathscr{O} = \ddot{\mathscr{O}}'_{\psi(N)}$ subject to the usual restrictions. We write

$$\ddot{\mathscr{O}}'_{\psi(N)} = \{\eta, 1_1, 1_2, \ldots, 1_{\psi(N)}, 2_1, \ldots, 2_{\psi(N)}, \ldots, K_1, \ldots, K_{\psi(N)}\},$$

where o_l refers to the l-th copy of object o. Second, copies of the objects are indistinguishable. Thus, nature draws true values and types as in the baseline

version (of the model) and players' expected true values and payoffs are formed as in the baseline version – i.e., using equation (2). The true values, expected true values, and payoffs are then extended from $\mathscr{O}'$ to $\ddot{\mathscr{O}}'_{\psi(N)}$ by treating the copies of each object identically. Thus, when $v_i^\dagger = (v_{\eta i}^\dagger, v_{1i}^\dagger, \ldots, v_{Ki}^\dagger)$ is player i's expected true values after learning her type θ_i, the copies of object o, i.e., $o_1, \ldots, o_{\psi(N)}$, each have an expected true value of $v_{oi}^\dagger$. Consequently, (i) player i's payoff is $u_i(o_l|\theta_i) = v_{oi}^\dagger$ for each object $o \in \mathscr{O}$ and each copy $l \in \{1, \ldots, \psi(N)\}$ and (ii) her payoff to η is $u_i(\eta|\theta_i) = v_{\eta i}^\dagger$. Given these payoffs, concepts like interim efficiency and ε-Bayesian incentive compatibility extend naturally.

We also extend the Simple Mechanism to allow for object replication by making it aware of these payoff extensions. Specifically, for each player i, after employing equation (4) and estimating $\hat{u}_i$ on $\mathscr{O}'$ using i's report $r_i \in \Theta$, it extends $\hat{u}_i$ to $\ddot{\mathscr{O}}'_{\psi(N)}$ by setting $\hat{u}_i(o_l|\theta_i) = \hat{u}_i(o|\theta_i)$ for each $o \in \mathscr{O}$ and each copy $l \in \{1, \ldots, \psi(N)\}$. It then computes the set of pure assignments that maximize the sum of the estimated payoffs and randomizes in the usual fashion.

Consider a version of the baseline model, which we call the **T-replication case**, where T denotes the number of replications. The case is defined via the primitives

$$\ddot{\mathscr{O}}_T = \cup_{l=1}^T \mathscr{O} = \{1_1, 1_2, \ldots, 1_T, 2_1, \ldots, 2_T, \ldots, K_1, \ldots, K_T\},$$

$\ddot{\mathscr{V}} = \mathscr{V}_\eta \times_{o \in \mathscr{O}} \ddot{\mathscr{V}}_o(T)$ where $\ddot{\mathscr{V}}_o(T) = \{(x_1, \ldots, x_T) \in \mathbb{R}^T | \exists v \in \mathscr{V}_o \text{ s.t. } x_1 = x_2 = \cdots = v\}$ for each $o \in \mathscr{O}$ (per the finite grid assumption),

$$f_{\ddot{v}}(v_\eta, v_{1_1}, \ldots, v_{2_1}, \ldots, v_{K_1}, \ldots, v_{K_T}) = f_v(v_\eta, v_{1_1}, v_{2_1}, \ldots, v_{K_1})$$

for each $(v_\eta, v_{1_1}, \ldots, v_{2_1}, \ldots, v_{K_1}, \ldots, v_{K_T}) \in \ddot{\mathscr{V}}$, α, and N. In this case, players' true values are defined directly on $\ddot{\mathscr{O}}'_T = \{\eta\} \cup \ddot{\mathscr{O}}_T$ (by treating copies of objects identically), rather than on $\mathscr{O}$ and then extended.[42] It is clear that players may have the same true values in the T-replication case and the replication extension when $T = \psi(N)$.

When $T = \psi(N)$, there is a natural bijection between the type-spaces of these environments: (i) a player's type is $(v_\eta, v_1, v_2, \ldots, v_K)$ in the replication

[42] We give the replication extension to avoid the technical challenges of modifying players types to account for growth in the number of copies in the T-replication case.

extension if and only if her type in the T-replication case is

$$(\eta, \underbrace{v_1, \ldots, v_1}_{T \text{ copies}}, \underbrace{v_2, \ldots, v_2}_{T \text{ copies}}, \ldots, \underbrace{v_K, \ldots, v_K}_{T \text{ copies}})$$

and (ii) a player's type is $\preceq$ in the replication extension if and only if her type in the T-replication case is $\preceq'$, where $\preceq$ and $\preceq'$ are such that, for any $o, o' \in \mathscr{O}'$ and any $l, l' \in \{1, \ldots, T\}$, we have (i) $o \preceq o'$ if and only if $o_l \preceq' o'_{l'}$ and (ii) $o \not\preceq o'$ if and only if $o_l \not\preceq' o'_{l'}$. Let $h : \Theta \to \ddot{\Theta}$ denote this bijection, where Θ denotes the replication extension type-space and $\ddot{\Theta}$ denotes the T-replication case type-space. It is clear that player i with type θ_i has the same expected true values and payoffs in the replication extension as she does in the T-replication case when her type is $h(\theta_i)$ and vice versa. (Also, let $h(\theta_1, \theta_2, \ldots, \theta_N) = (h(\theta_1), h(\theta_2), \ldots, h(\theta_N))$ and $h(r_1, r_2, \ldots, r_N) = (h(r_1), \ldots, h(r_N))$ for any $(\theta_1, \ldots, \theta_N)$ or $(r_1, \ldots, r_N)$ in Θ^N.)

The payoff equivalence implies that if player i, with type θ_i, makes report r_i to the Simple Mechanism in the replication extension, then she has the same payoff to the mechanism as in the T-replication case when she has type $h(\theta_i)$ and makes report $h(r_i)$ to the mechanism. In symbols, $U_i^{M_S}(r_i|\theta_i) = \ddot{U}_i^{M_S}(h(r_i)|h(\theta_i))$ for all $r, \theta_i \in \Theta$ and each player i when $T = \psi(N)$, where $U_i^{M_S}(r_i|\theta_i)$ denotes player i's payoff, given her type $\theta_i \in \Theta$, to making report $r_i \in \Theta$ to the Simple Mechanism in the replication extension and (in an abuse of notation) $\ddot{U}_i^{M_S}(r'_i|\theta'_i)$ denotes i's payoff, given her type $\theta'_i \in \ddot{\Theta}$, to making report $r'_i \in \ddot{\Theta}$ to the Simple Mechanism in the T-replication case.

To better understand this equivalence, suppose an arbitrary player k reports r_k to the Simple Mechanism in the replication extension and $h(r_k)$ in the T-replication case. If these reports were her types, it is evident that her expected true values would be the same in both environments. Thus, the Simple Mechanism estimates she has the same payoff in both environments. It follows that, for any joint vector of players' reports r in the replication extension, the mechanism has the same objective function and thus randomizes over the same pure assignments as in the T-replication case when players report $h(r)$. Given this, given the payoff equivalence, and given the fact that each joint vector of types θ in the replication extension has the same probability as the joint type vector $h(\theta)$ in the T-replication case (and vice versa), it is straightforward to show that $U_i^{M_S}(r_i|\theta_i) = \ddot{U}_i^{M_S}(h(r_i)|h(\theta_i))$ for each player i

and each $r_i, \theta_i \in \Theta$.[43]

Proof of Proposition 4. To establish the first part of the proposition, it suffices to provide an example where every interim efficient and symmetric mechanic is not ε-Bayesian incentive compatible. We do this in Example A (below), while further elaborating on the intuition for the failure of ε-Bayesian incentive compatibility.

Fix $\varepsilon > 0$ and consider player i. To establish the second part of the proposition, it suffices to show that there are (i) a linking function $\bar{\psi}$ and (ii) a $\bar{N}$ such that $\max_{r_i \in \Theta} U_i^{M_S}(r_i|\theta_i) \leq U_i^{M_S}(\theta_i|\theta_i) + \varepsilon$ for all $N \geq \bar{N}$ for each $\theta_i \in \Theta$ in the replication extension. (This is sufficient because all players with the same type have the same payoff.)

To do this, first consider the T-replication case. Since this case is a version of the baseline model, Proposition 3 holds and gives there is some N_T such that $\max_{r_i \in \ddot{\Theta}} \dot{U}_i^{M_S}(r_i|\theta_i) \leq \dot{U}_i^{M_S}(\theta_i|\theta_i) + \varepsilon$ for all $N \geq N_T$ for each $\theta_i \in \ddot{\Theta}$. Thus, there is a sequence of thresholds $N_1, N_2, \ldots, N_T, \ldots$ that ensure ε-Bayesian incentive compatibility in the $1, 2, \ldots, T, \ldots$ replication cases respectively. It is without loss to suppose $N_1 < N_2 < \cdots < N_T < \cdots$. We use this sequence of thresholds to construct $\bar{\psi}$ as follows,

$$\bar{\psi}(n) = \begin{cases} 1 & \text{if } n < N_2 \\ 2 & \text{if } N_2 < n \leq N_3 \\ \vdots & \vdots \\ T & \text{if } N_T < N \leq N_{T+1} \\ \vdots & \vdots \end{cases}.$$

It is evident that this function is weakly monotone on $\mathbb{N}$ and strictly monotone on an infinite sub-set of $\mathbb{N}$, implying $\lim_{n\to\infty} \bar{\psi}(n) = \infty$.

Returning to the replication extension, when the number of copies of each object is determined by $\bar{\psi}(N)$, we claim that $\bar{N} = N_1$ is the required threshold. To see this, pick some number of players $N \geq \bar{N}$ and observe that $N \in [N_{T'}, N_{T'+1})$ for some $T' \geq 1$. Since (i) $\max_{r_i \in \ddot{\Theta}} \dot{U}_i^{M_S}(r_i|\theta_i) \leq \dot{U}_i^{M_S}(\theta_i|\theta_i) + \varepsilon$ for each $\theta_i \in \ddot{\Theta}$ in the T'-replication case and since (ii)

$$U_i^{M_S}(r_i|\theta_i) = U_i^{M_S}(h(r_i)|h(\theta_i))$$

[43] We formally establish this equality in the Online Appendix.

for all $r_i, \theta_i \in \Theta$ because $\bar{\psi}(N) = T'$, we have, $\max_{r_i \in \Theta} U_i^{M_S}(r_i|\theta_i) =$

$$\max_{r_i \in \Theta} U_i^{M_S}(h(r_i)|h(\theta_i)) \leq \ddot{U}_i^{M_S}(h(\theta_i)|h(\theta_i)) + \varepsilon = U_i^{M_S}(\theta_i|\theta_i) + \varepsilon,$$

for each $\theta_i \in \Theta$ in the replication extension. □

Example A. Failure of ε-Bayesian Incentive Compatibility with an Arbitrary Linking Function.

Suppose there are $N \geq 2$ players and there is a single object o with $\psi(N) = \lfloor N/2 \rfloor$ copies. Further, suppose $\mathscr{V} = \{0\} \times \{\frac{1}{2}, 1\}$, with $f_v(\cdot) = \frac{1}{2}$ and $\alpha = 1$. That is, players observe their values and there are two equally likely types of players: type "As" who have values $(v_\eta, v_o) = (0, 1)$ and type Bs who have values $(v_\eta, v_o) = (0, \frac{1}{2})$.

After receiving reports, it is evident that any interim efficient mechanism assigns the object to type A reporters before assigning any remaining copies to type B reporters. Furthermore, it is readily verified that any symmetric mechanism endows each type A reporter with the same probability of receiving a copy of the object, likewise for each type B reporter. Let M denote an interim efficient and symmetric mechanism.

To show that there is no symmetric, interim efficient, and ε-Bayesian incentive compatible mechanism in this example, focus on the truth-telling incentives player for 1 when she is a type B. Let A_l be the event there are $l \in \{0, 1, \ldots, N-1\}$ other players who are type A. If all other players report truthfully, then given l the probability a type A gets an object is $\frac{\min\{\lfloor N/2 \rfloor, l\}}{l}$ and the probability a type B gets an object is $\frac{\max\{\lfloor N/2 \rfloor - l, 0\}}{N-l}$.[44] Thus, player 1's payoff to telling the truth to M is

$$U_1^M(B|B) = \frac{1}{2} \sum_{l=0}^{N-1} \frac{\max\{\lfloor N/2 \rfloor - l, 0\}}{N-l} \binom{N-1}{l} \left(\frac{1}{2}\right)^{N-1},$$

where $\binom{N-1}{l} \left(\frac{1}{2}\right)^{N-1}$ is the probability of l type A players and $N-l$ type B players. Alternatively, if player 1 were to lie and report she is a type A, she

[44] When there are l type As and $\lfloor N/2 \rfloor$ objects, either (i) every type A gets an object with certainty or (ii) they enter into a uniform lottery for these objects (by symmetry). After the type As are assigned, there are $\max\{\lfloor N/2 \rfloor - l, 0\}$ objects for the type Bs, so either (i') every type B get an object with certainty or (ii') the type Bs enter a uniform lottery for the remaining objects. Thus, $\frac{\min\{\lfloor N/2 \rfloor, l\}}{l}$ is the probability a type A is assigned the object and $\frac{\max\{\lfloor N/2 \rfloor - l, 0\}}{N-l}$ is the probability a type B is assigned an object (since player 1 is a type B).

would get the object with probability $\frac{\min\{\lfloor N/2\rfloor, l+1\}}{l+1}$ given l. Thus, her payoff to lying to M is

$$U_1^M(A|B) = \frac{1}{2}\sum_{l=0}^{N-1} \frac{\min\{\lfloor N/2\rfloor, l+1\}}{l+1}\binom{N-1}{l}\left(\frac{1}{2}\right)^{N-1}.$$

Hence, player 1's gain to lying is

$$\begin{aligned}
&U_1^M(A|B) - U_1^M(B|B) = \\
&\frac{1}{2}\sum_{l=0}^{N-1}\left(\frac{\min\{\lfloor N/2\rfloor, l+1\}}{l+1} - \frac{\max\{\lfloor N/2\rfloor - l, 0\}}{N-l}\right)\binom{N-1}{l}\left(\frac{1}{2}\right)^{N-1} = \\
&\qquad \frac{1}{2}\sum_{l=0}^{\lfloor N/2\rfloor - 1}\left(1 - \frac{\lfloor N/2\rfloor - l}{N-l}\right)\binom{N-1}{l}\left(\frac{1}{2}\right)^{N-1} + \\
&\qquad\qquad \frac{1}{2}\sum_{l=\lfloor N/2\rfloor}^{N-1}\left(\frac{\lfloor N/2\rfloor}{l+1}\right)\binom{N-1}{l}\left(\frac{1}{2}\right)^{N-1}.
\end{aligned}$$

It is clear that player 1 always gains from lying; this is illustrated in Figure A.1, which plots the value of this difference for N from 2 to 500. In fact, algebra shows (i) that the difference is always at least $\frac{1}{4}$ and (ii) that $\lim_{N\to\infty} U_1^M(A|B) - U_1^M(B|B) = \frac{1}{2}$. In other words, not only does player 1 gain a strictly positive, minimum amount from lying, but her gain actually increases with the number of players! It follows that, any symmetric and interim efficient mechanism M is not ε-Bayesian incentive compatible. The intuition for this outcome is that, with high probability, there are approximately as many objects as there are type A players. Thus, any symmetric and interim efficient mechanism typically assigns (almost) all of the objects to the type As, leaving few to none for the type Bs. Thus, a type B can obtain the object with high probability by lying. All of this is possible because replication negates competition for the object among the type As. If competition were present among the type As, then a type B's gains to lying would evaporate and ε-Bayesian incentive compatibility would obtain (since no type A has incentive to lie). $\triangle$

Example B. ε-Bayesian Incentive Compatibility when Replication Occurs Slowly.

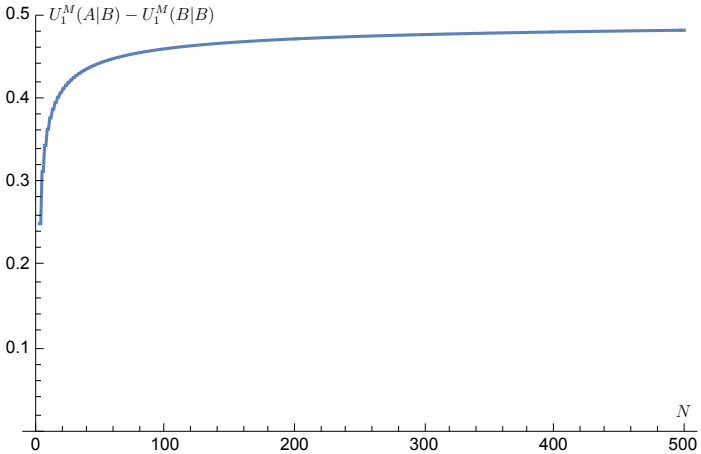

Figure 1: Example A, $U_1^M(A|B) - U_1^M(B|B)$ for $N \in \{2, 3, \ldots, 500\}$

Consider the same setup as Example A, but instead suppose that $\psi(N) = \lfloor \ln(N)/2 \rfloor$ and $N \geq 8$.[45] Then, algebra shows that player 1's gain to lying to the Simple Mechanism is

$$U_1^{M_S}(A|B) - U_1^{M_S}(B|B) =$$
$$\frac{1}{2}\sum_{l=0}^{N-1}\left(\frac{\min\{\lfloor \ln(N)/2 \rfloor, l+1\}}{l+1} - \frac{\max\{\lfloor \ln(N)/2 \rfloor - l, 0\}}{N-l}\right)\binom{N-1}{l}\left(\frac{1}{2}\right)^{N-1}.$$

While this difference is initially positive, it is evident that it shrinks to zero as $N \to \infty$. This behavior is illustrated in Figure A.2 which plots the gain to lying for N from 8 to 500. The intuition is that, unlike in Example A, the number of objects increases sufficiently slowly so that competition intensifies among the type As, limiting player 1's and thus any type B's gain to misrepresentation. Since no type A player has incentive to lie, ε-Bayesian incentive compatibility obtains. $\triangle$

References

Abdulkadiroglu, A., Che, Y., & Yasuda, Y. (2008). Expanding "Choice" in School Choice. Working Paper, Duke University.

[45] We take $N \geq 8$ so that $\psi(N) \geq \lfloor \ln(8)/2 \rfloor = 1$.

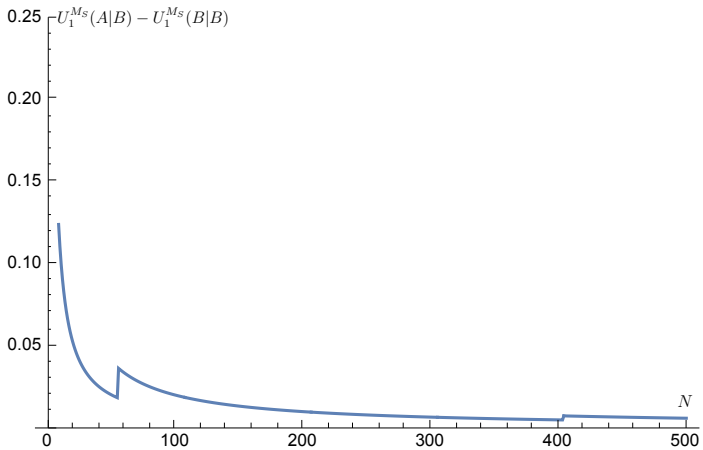

Figure 2: Example B, $U_1^{M_S}(A|B) - U_1^{M_S}(B|B)$ for $N \in \{8, 9, \ldots, 500\}$

Abdulkadiroglu, A., Che, Y., & Yasuda, Y. (2011, February). Resolving Conflicting Preferences in School Choice: The "Boston Mechanism" Reconsidered. The American Economic Review, 101(1), 399–410.

Abdulkadiroglu, A., & Sonmez, T. (1998). Random Serial Dictatorship and the Core From Random Endowments in House Allocation Problems. Econometrica: Journal of the Econometric Society, 66(3), 689–701.

Abdulkadiroglu, A., & Sonmez, T. (2003). School Choice: A Mechanism Design Approach. American Economic Review, 93(3), 729-747.

Azevedo, E., & Budish, E. (2018). Strategy proofness in the large. Working Paper, University of Pennsylvania.

Bogomolnaia, A., & Moulin, H. (2001). A New Solution to the Random Assignment Problem. Journal of Economic Theory, 100(2), 295–328.

Bu, N. (2014). Characterizations of the Sequential Priority Rules in the Assignment of Object Types. Social Choice and Welfare, 43(635-645).

Budish, E. (2011, December). The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes. Journal of Political Economy, 119(6), 1061–1103.

Budish, E., Che, Y., Kojima, F., & Milgrom, P. (2013). Designing Random Allocation Mechanisms: Theory and Applications. American Economic Review, 103(2), 585-623.

Che, Y., & Tercieux, O. (2013). Efficiency and Stability in Large Matching Markets. Working Paper, Columbia University.

Ehlers, L., Majumdar, D., Mishra, D., & Sen, A. (2014). Continuous Cardinal

Incentive Compatible Mechanisms are Ordinal. Working Paper, University of Montreal.

Feller, W. (1968). An Introduction to Probability Theory and its Applications. John Wiley and Sons.

Gale, D., & Shapley, L. (1962). College Admissions and the Stability of Marriage. The American Mathematical Monthly, 69(1), 9–15.

Hafalir, I., & Miralles, A. (2014). Welfare Maximizing Assignment of Agents to Hierarchical Positions. Working Paper, Carnegie Mellon University.

He, Y., Miralles, A., Pycia, M., & Yan, J. (2018). A Pseudo-Market Approach to Allocation with Priorities. American Economic Journal: Microeconomics, 10(3), 272-314.

Hylland, A., & Zeckhauser, R. (1979). The Efficient Allocation of Individuals to Positions. Journal of Political Economy, 87(2), 293-314.

Kasajima, Y. (2013). Probabilistic Assignment of Indivisible Goods with Single-Peaked Preferences. Social Choice and Welfare, 41, 203-215.

Liu, Q., & Pycia, M. (2016). Ordinal Efficiency, Fairness, and Incentives in Large Markets. Working Paper, UC Los Angeles.

Miralles, A. (2008). School Choice: The Case for the Boston Mechanism. Working Paper, Autonomous University of Barcelona.

Miralles, A., & Pycia, M. (2014). Cardinal Efficiency in Assignments Without Transfers. Working Paper, UC Los Angeles.

Nguyen, T., Peivandi, A., & Vohra, R. (2014). One-Sided Matching with Limited Complementarities. Working Paper, University of Pennsylvania.

Nicolo, A., & Rodriguez-Alvarez, C. (2013). Incentive Compatibility and Feasibility Constraints in Housing Markets. Social Choice and Welfare, 41, 625-635.

Pycia, M. (2011, June). The Cost of Ordinality. Working Paper, UC Los Angeles, 1–4.

Roth, A., Sonmez, T., & U., U. (2004). Kidney Exchange. Quarterly Journal of Economics, 119(2), 457-488.

Roth, A., & Sotomayor, M. (1990). Two-sided matching: A study in game-theoretic modeling and analysis. Cambridge University Press.

Shapley, L., & Scarf, H. (1974). On Cores and Indivisibility. Journal of Mathematical Economics, 1(1), 23–37.

Sonmez, T., & Unver, U. (2010). House Allocation with Existing Tenants: A Characterization. Games and Economic Behavior, 69, 425-445.

Zhou, L. (1990). On a Conjecture by Gale About One-Sided Matching Problems. Journal of Economic Theory, 52(1), 123–135.

Journal of Mechanism and Institution Design
ISSN: 2399-844X (Print), 2399-8458 (Online)
DOI:10.22574/jmid.2018.12.004

MECHANISMS FOR HOUSE ALLOCATION WITH EXISTING TENANTS UNDER DICHOTOMOUS PREFERENCES

Haris Aziz

Data61, CSIRO and UNSW Sydney, Australia
haris.aziz@data61.csiro.au

ABSTRACT

We consider house allocation with existing tenants in which each agent has dichotomous preferences. We present strategyproof, polynomial-time, and (strongly) individually rational algorithms that satisfy the maximum number of agents. For the endowment only model, one of the algorithms also returns a core-stable allocation.

Keywords: House allocation, core, dichotomous preferences.

JEL Classification Numbers: C62, C63, C78.

1. INTRODUCTION

Principled allocation of resources is a central problem faced by society. We consider an allocation setting in which there is a set of agents $N = \{1, \ldots, n\}$, a set of houses $H = \{h_1, \ldots, h_m\}$, each agent owns at most one and possibly zero houses from H and there may be houses that are not owned by any agent. In the literature the setting is termed as *house allocation with existing tenants* (Abdulkadiroğlu & Sönmez, 1999; Sönmez & Ünver, 2010). If each agent owns exactly one house and each house is owned by some agent, the setting is equivalent to the *housing market setting* (D. J. Abraham et al., 2005; Aziz & de Keijzer, 2012; Ma, 1994; Plaxton, 2013; Shapley & Scarf, 1974).

The author thanks the reviewers for helpful feedback which helped improve the presentation. The author also thanks Péter Biró, Greg Plaxton, Zhaohong Sun, and the anonymous reviewers for useful comments and pointers.

If no agent owns a house, the setting is equivalent to the house allocation model (Hylland & Zeckhauser, 1979; Svensson, 1994).

The setting captures various allocations settings including allocation of dormitory rooms as well as kidney exchange markets. We consider housing market with existing tenants setting with the assumption that each agent has dichotomous preferences with utility zero or one.

Dichotomous preferences are prevalent in many settings. A kidney may either be compatible or incompatible for a patient. A dormitory room may be within budget or outside budget for a student. Bogomolnaia & Moulin (2004) give numerous other examples where dichotomous preferences make sense in allocation and matching problems. Interestingly, since dichotomous preferences require indifference in preferences, they are not covered by many models considered in the literature that assume strict preferences (Abdulkadiroğlu & Sönmez, 1999; Sönmez & Ünver, 2010). If a house is acceptable, we will refer to it as an acceptable house. If an agent gets a house that is acceptable, we will say that the agent is *satisfied*.

For the setting, we study how to allocate the houses to the agents in a desirable manner. There are several properties that capture how to allocate resources in a desirable manner. We consider the standard normative properties in market design: (i) *Pareto optimality*: there should be no other allocation in which each agent is at least as happy and at least one agent is strictly happier (ii) *individual rationality (IR)*: no agent should have an incentive to leave the allocation program (iii) *strategyproofness*: no agent should have an incentive to misreport his preferences. We also consider a stronger notion of individual rationality that requires that either an agent keeps his endowment or gets some house that is strictly better. Finally, we consider a property stronger than Pareto optimality: maximizing the number of agents who are satisfied. If we rephrase dichotomous preferences in terms of 1-0 utilities, the goal is equivalent to maximizing social welfare. Under 1-0 utilities, an allocation maximizes welfare if a maximum number of agents get utility 1 or in other words a maximum number of agents are satisfied. Throughout the paper, when we refer to (utilitarian) welfare, we will assume that the underlying cardinal utilities are 1-0.

The main results are as follows:

Theorem 1. *For house allocation with existing tenants with dichotomous preferences, there exists an algorithm (MSIR) that is polynomial-time, strategyproof, and allocates acceptable houses to a maximum number of agents*

subject to strong individual rationality. In the endowment setting in which each house is initially owned by an agent, the algorithm is core stable.

Theorem 2. *For house allocation with existing tenants with dichotomous preferences, there exists an algorithm (MIR) that is polynomial-time, strategyproof, allocates acceptable houses to a maximum number of agents, and is Pareto optimal.*

2. RELATED WORK

House allocation with existing tenants model was introduced by Abdulkadiroğlu & Sönmez (1999). They assumed that agents have strict preferences and introduced mechanisms that are individually rational, Pareto optimal and strategyproof. The results do not apply to the setting with dichotomous preferences.

In the restricted model of housing markets, Jaramillo & Manjunath (2012) proposed a mechanism called TCR that is polynomial-time, IR (but not S-IR), Pareto optimal and strategyproof even if agents have indifferences. The results also apply to housing markets under dichotomous preferences. The MIR and MSIR mechanisms are relatively simpler and the arguments for their properties are simple and direct as well.

Another natural approach to allocation problems is to satisfy the maximum number of agents by computing a maximum weight matching. However, such an approach violates individual rationality. Krysta et al. (2014) proposed algorithms for house allocation with 1-0 utility. However their algorithms are for the model in which agents do not have any endowment.

For kidney exchange with 1-0 utilities, D. Abraham et al. (2007) and Biro et al. (2009) presented an algorithm that is strongly individually rational and satisfies the maximum number of agents. The algorithm does not cater for the cases where agents own acceptable houses, or there may be houses that are not owned by any agents. Furthermore, they did not establish strategyproofness or core stability of their mechanism.

3. PRELIMINARIES

Let $N = \{1, \ldots, n\}$ be a set of n agents and $H = \{h_1, \ldots, h_m\}$ a set of m houses. The endowment function $e : N \to H \cup \{\text{null}\}$ assigns to each agent the house he originally owns. The value $e(i)$ is null if agent i does not own a house.

Each agent has complete and transitive preferences $\succsim_i$ over the houses and $\succsim = (\succsim_1, \dots \succsim_n)$ is the preference profile of the agents. The *housing market* is a quadruple $M = (N, H, e, \succsim)$. For $S \subseteq N$, we denote $\bigcup_{i \in S}\{e(i)\}$ by $e(S)$. A function $x : S \to H$ is an *allocation* on $S \subseteq N$ tha allocates house $x(i)$ to agent $i \in N$. The goal is to allocate the houses in a mutually beneficial and efficient way.

Since we consider dichotomous preferences, we will denote by A_i the set of houses acceptable by $i \in N$. These houses are *acceptable* to i whereas the other houses are unacceptable to i. Agent i is *satisfied* when he gets an acceptable house. Since we consider welfare as well, we will assume that acceptable houses give utility one to the agent and unacceptable houses give utility zero. Hence getting an unacceptable house is equivalent to not getting a house.

An allocation x is *individually rational (IR)* if $x(i) \succsim_i e(i)$ for all $i \in N$. IR requires that an agent endowed with an acceptable house is allocated an acceptable house. An allocation x is *strongly individually rational (S-IR)* if $x(i) = e(i)$ or $x(i) \succ_i (e(i))$ for all $i \in N$. SIR requires that an agent only changes his house if he gets a strictly more preferred house.

A coalition $S \subseteq N$ *blocks* an allocation x on N if there exists an allocation y on S such that for all $i \in S$, $y(i) \in e(S)$ and $y(i) \succ_i x(i)$. An allocation x on N is in the *core (C)* of market M if it admits no blocking coalition. An allocation that is in the core is also said to be *core stable*. A coalition $S \subseteq N$ *weakly blocks* an allocation x on N if there exists an allocation y on S such that for all $i \in S$, $y(i) \in e(S)$, $y(i) \succsim_i x(i)$, and there exists an $i \in S$ such that $y(i) \succ_i x(i)$. An allocation x on N is in the *strict core (SC)* of market M if it admits no weakly blocking coalition. An allocation that is in the strict core is also said to be *strict core stable*. An allocation is *Pareto optimal (PO)* if N is not a weakly blocking coalition. It is clear that strict core implies core and also Pareto optimality. Core implies individual rationality.

From now on we will assume 1-0 utilities for all the statements.

Example 1. *Consider a house allocation with existing tenants setting:*

- $N = \{1,2,3,4,5\}$;
- $H = \{h_1, h_2, h_3, h_4\}$;
- $e(i) = h_i$ *for* $i \in \{1, \dots, 4\}$; $e(5) = null$
- $A_1 = \{h_2\}$; $A_2 = \{h_3\}$; $A_3 = \{h_1\}$, $A_4 = \{h_3\}$, $A_5 = \{h_4\}$

A feasible S-IR outcome is allocation x such that $x(1) = h_2$, $x(2) = h_3$, $x(3) = h_1$, $x(4) = h_4$, and $x(5) = null$.

A feasible IR and Pareto optimal outcome is allocation y such that $y(1) = h_2$, $y(2) = h_3$, $y(3) = h_1$, $y(4) = null$, and $y(5) = h_4$.

It is obvious that S-IR implies IR. We also observe the following.

Proposition 1. *An allocation that maximizes welfare subject to S-IR may have less welfare than an allocation that maximizes welfare subject to IR.*

Proof. Consider the setting in which there are two agents and both agents have zero value houses but one agent wants the other agent's house. The only feasible S-IR allocation is the endowment allocation whereas there exists an IR allocation in which one agents gets an acceptable. □

4. MECHANISMS

We present two mechanisms: MSIR and MIR. MSIR satisfies the maximum number of agents subject to the S-IR constraint. MIR satisfies the maximum number of agents subject to the IR constraint. Both mechanisms are based on constructing a bipartite graph that admits a perfect matching and repeatedly modifying the graph while still ensuring that the graph still has a maximum weight perfect matching. The mechanisms are parametrized by a permutation π that is a function $\pi : N \to N$. The function specifies an ordering of the agents in N so that $\pi(j)$ is the j-th agent in the ordering.

4.1. The MSIR Mechanism

The MSIR mechanism is specified as Algorithm 1.

Example 2. *Let us illustrate how the MSIR mechanism works on the setting in Example 1. Suppose $\pi = 12345$. First we form a graph in Figure 1. We find a perfect matching that satisfies the most number of agents subject to S-IR. Such a matching can satisfy at most 3 agents. We go through permutation 12345 and check whether there exists a matching that satisfies 3 agents, does not violate S-IR, satisfied the guarantees of previous agents in the permutation, and still satisfies the current agent in the permutation. If yes, that agent is guaranteed to be satisfied from now. After agents 1, 2, and 3 are processed, we know that 1 should get h_2, 2 should get h_3 and 3 should get h_1. S-IR requires that 4 and*

Algorithm 1 MSIR

Input: $(N, H, (A_1, \ldots, A_n), e)$
Output: Allocation $(x(1), \ldots, x(n))$

1 Consider weighted graph $G = (A \cup B, E, w)$. Set k to $|n-m|$. If $m > n$, then $A = N \cup D_N$ where $D_N = \{d_1, \ldots, d_k\}$. If $n > m$, then $B = H \cup D_H$ where $D_H = \{o_1, \ldots, o_k\}$. If $n = m$, then $A = N$ and $B = H$.
E is defined as follows:

- For each $a \in N$ and $b = e(a)$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 0$ if $b \notin A_a$ and $w(\{a,b\}) = 1$ if $b \in A_a$.
- For each $a \in N$ such that $e(a) \neq null$, $e(a) \notin A_a$, and each $b = H \setminus \{e(a)\}$ such that $b \in A_a$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 1$.
- For each $a \in N$ such that $e(a) = null$ and each $b \in B$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 0$ if $b \notin A_a$ and $w(\{a,b\}) = 1$ if $b \in A_a$.
- For each $b \in D_H$ and each $a \in \{a \in N : e(a) = null\}$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 0$.
- For each $a \in D_N$ and each $b \in B$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 0$.

{Note that if an agent has an acceptable endowment, then he only has an edge to his endowment. An agent with an unacceptable endowment only has edges with his endowment as well as the acceptable houses.}

2 Compute a maximum weight perfect matching of G. Let the weight of the matching be W.
3 Take a permutation (ordering) π of the elements in N. Denote by $\pi(i)$ the i-th element in the ordering.
4 **for** i=1 to n **do**
5 Set $t_{\pi(i)}$ to 1.
6 Remove from G each edge $\{\pi(i), b\} \in E$ such that $w(\{\pi(i), b)\}) = 0$. Compute a maximum weight perfect matching of G. If the weight is less than W or if there does not exist a perfect matching, then put back the removed edges and set $t_{\pi(i)}$ to 0. {Variable $t_{\pi(i)}$ keeps track of whether agent $\pi(i)$ should get an acceptable house or not.}
7 **end for**
8 Compute a maximum weight perfect matching M of G. Consider the allocation x in which each agent $i \in N$ gets a house that it is matched to in M. If $i \in N$ is matched to a dummy house, its allocation is null.
9 **return** $(x(1), \ldots, x(n))$

5 do not get any new house. Hence the result is the same as allocation x in Example 1.

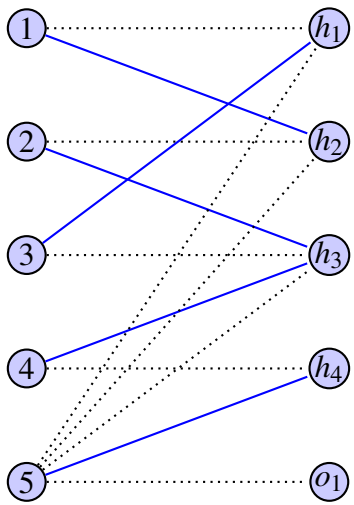

Figure 1: Initial graph G when running MSIR over the instance in Example 1. The solid edges have weight 1 whereas the dotted edges have weight 0. A perfect matching with the maximum weight has weight 3.

Proposition 2. *MSIR runs in polynomial-time.*

Proof. The graph G has $O(n+m)$ vertices. The graph is modified at most $O(n)$ times. Each time, a maximum weight perfect matching is computed that takes polynomial time $|V(G)|^3$ (Korte & Vygen, 1999). □

Proposition 3. *MSIR returns an allocation that is S-IR.*

Proof. Throughout the algorithm, we make sure that G admits a perfect matching. If a modification to G leads to a lack of a perfect matching, then such a modification is reversed. An agent with an endowment only has an edge to his endowment or to an acceptable house. Therefore while a perfect matching exists, an agent with a non-null endowment cannot be matched to an unacceptable house other than his endowment. Thus MIR returns an allocation that is S-IR. □

Proposition 4. *MSIR returns an allocation that satisfies the maximum number of agents subject to S-IR.*

Proof. Throughout the algorithm, we make sure that G admits a perfect matching which ensures that the corresponding allocation satisfies S-IR. An agent

with an acceptable endowment does not have an edge to any other house so he has to be allocated his endowment. An agent with an unacceptable endowment either has an edge to his endowment and to houses that are acceptable. Hence, in a perfect matching, the agent either gets his endowment or an acceptable house. Given this condition, we compute the maximum weight perfect matching. This implies that the corresponding allocation satisfies the maximum number of agents under the S-IR constraint. □

Corollary 1. *MSIR returns an allocation that is Pareto optimal among the set of S-IR allocations.*

Proof. Assume for contradiction that the MSIR allocation is Pareto dominated by an S-IR allocation. But this means that the MSIR allocation does not satisfy the maximum number of agents. □

Proposition 5. *MSIR is strategyproof.*

Proof. MSIR returns a perfect matching of weight W. During the running of MSIR, each time a modification is made to the graph G, it is ensured that G admits a perfect matching of weight W. Assume for contradiction that MSIR is not strategyproof and some agent $i \in N$ with turn k in permutation π gets a more preferred house when he misreports. This means that agent i gets an unacceptable house when he reports the truthful preference A_i. Let the allocation be x. This implies that in permutation π, when i's turn comes, there exists no feasible maximum weight perfect matching of weight W in which i gets an acceptable house and each agent $\pi(j)$ preceding i in permutation π gets $t_{\pi(j)}$ acceptable houses. Since i can get a more preferred house by misreporting, i gets an acceptable house if he reports A'_i. Let such an allocation be x'. Note that x' is a feasible maximum weight perfect matching even when i tells the truth and even if each agent $\pi(j)$ preceding i in permutation π gets $t_{\pi(j)}$ acceptable houses. But this is a contradiction because there does exist a feasible maximum weight perfect matching of weight W in which i gets an acceptable house and each agent preceding i in permutation π gets $t_{\pi(j)}$ acceptable houses. □

Lemma 1. *For 1-0 utilities, any S-IR welfare maximizing allocation is core stable.*

Proof. Suppose that there is a blocking coalition S. It can only consist of agents who did not get an acceptable house in the allocation. S-IR implies that

agents who originally own an acceptable house keep the acceptable house. If an agent $i \in N$ is in S who does not originally own any house, he cannot be part of S because he has nothing to give to other agents, so other agents can satisfy each other without letting i be a member of S. Therefore, S consists of those agents who owned an unacceptable house and are allocated an unacceptable house. Due to S-IR, agents in S are allocated their own house. Now if S admits a blocking coalition this implies that the S-IR welfare maximizing allocation was not S-IR welfare maximizing which is a contradiction. □

In view of the lemma, we derive the following statement.

Proposition 6. *MSIR returns an allocation that is core stable.*

4.2. The MIR Mechanism

The MIR mechanism is specified as Algorithm 2. The main difference between MIR and MSIR is that we impose the less restrictive IR constraint for MIR rather than the S-IR constraint. In MIR, an agent who owns an acceptable house is willing to get some other acceptable house. If the agent owns an unacceptable house, he is willing to give that house to someone else even if he does not get an acceptable house in return.

Example 3. *Let us illustrate how the MIR mechanism works on the setting in Example 1. Suppose* $\pi = 12345$. *First we form a graph in Figure 2. We then find a perfect matching that satisfies the most number of agents subject to IR. Such a matching can satisfy at most 4 agents. We go through permutation* 12345 *and check whether there exists a matching that satisfies 4 agents, does not violate IR, satisfies the guarantees of the previous agents in the permutation, and still satisfies the current agent in the permutation. If yes, we require that agent to be satisfied. After agents 1, 2, and 3 are processed, we know that 1, 2, and 3 can be satisfied by giving them* h_2, h_3, *and* h_1 *respectively. Agent 4 cannot be satisfied while requiring agents 1, 2, and 3 to be satisfied. The last agent 5 can be satisfied without violating IR or resulting in some agent among 1, 2, and 3 to be not satisfied. The outcome is the same as allocation* y *in Example 1.*

Proposition 7. *MIR runs in polynomial-time.*

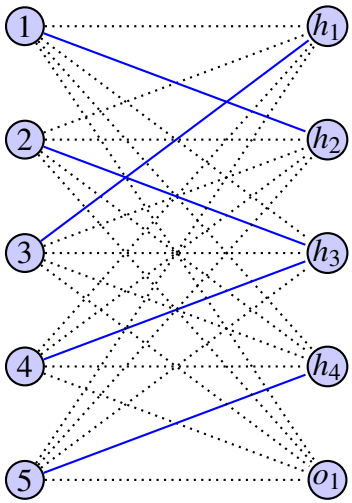

Figure 2: Initial graph G when running MIR over the instance in Example 1. The solid edges have weight 1 whereas the dotted edges have weight 0. A perfect matching with the maximum weight has weight 4.

Proof. The graph G has $O(n+m)$ vertices. The graph is modified at most $O(n)$ times. Each time, a maximum weight perfect matching is computed that takes time $|V(G)|^3$ (Korte & Vygen, 1999). □

Proposition 8. *MIR returns an allocation that is IR.*

Proof. Throughout the algorithm, we make sure that G admits a perfect matching. If a modification to G leads to a lack of a perfect matching, then such a modification is reversed. An agent with an acceptable endowment only has an edge to his endowment or to other acceptable houses. Therefore while a perfect matching exists, an agent with an acceptable endowment can only be matched to an acceptable house. Thus MIR returns an allocation that is IR. □

Proposition 9. *MIR returns an allocation that satisfies the maximum number of agents subject to IR.*

Proof. Throughout the algorithm, we make sure that G admits a perfect matching which ensures that the corresponding allocation satisfies IR. Given this condition, we compute the maximum weight perfect matching. This implies that the corresponding allocation satisfies the maximum number of agents under the IR constraint. □

Algorithm 2 MIR

Input: $(N, H, (A_1, \ldots, A_n), e)$
Output: $(x(1), \ldots, x(n))$

1 Consider weighted graph $G = (A \cup B, E, w)$. Set k to $|n-m|$. If $m > n$, then $A = N \cup D_N$ where $D_N = \{d_1, \ldots, d_k\}$. If $n > m$, then $B = H \cup D_H$ where $D_H = \{o_1, \ldots, o_k\}$. If $n = m$, then $A = N$ and $B = H$.
E is defined as follows:

- For each $a \in N$ and $b = e(a)$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 0$ if $b \notin A_a$ and $w(\{a,b\}) = 1$ if $b \in A_a$.
- For each $a \in N$ and each $b \in H \setminus \{e(a)\}$ such that $b \in A_a$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 1$.
- For each $a \in N$ such that $e(a) = null$ and each $b \in B$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 0$ if $b \notin A_a$ and $w(\{a,b\}) = 1$ if $b \in A_a$.
- For each $a \in N$ such that $e(a) = h$ and $h \notin A_a$ and each $b \in B$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 0$ if $b \notin A_a$ and $w(\{a,b\}) = 1$ if $b \in A_a$.
- For each $b \in D_H$ and each $a \in \{a \in N : e(a) \notin A_a\}$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 0$.
- For each $a \in D_N$ and each $b \in B$, add $\{a,b\} \in E$ with $w(\{a,b\}) = 0$.

{Note that if an agent has an acceptable endowment, he only has edges with acceptable houses. An agent with an unacceptable endowment has an edge to every $b \in B$.}

2 Compute a maximum weight perfect matching of G. Let the weight of the matching be W.
3 Take a permutation (ordering) π of the elements in N. Denote by $\pi(i)$ the i-th element in the ordering.
4 **for** i=1 to n **do**
5 Set $t_{\pi(i)}$ to 1.
6 Remove from G each edge $\{\pi(i), b\} \in E$ such that $w(\{\pi(i), b)\}) = 0$. Compute the maximum weight perfect matching of G. If the weight is less than W or if there does not exist a perfect matching, then put back the removed edges and set $t_{\pi(i)}$ to 0. {Variable $t_{\pi(i)}$ keeps track of whether agent $\pi(i)$ should get an acceptable house or not.}
7 **end for**
8 Compute a maximum weight perfect matching M of G. Consider the allocation x in which each agent $i \in N$ gets a house that it is matched to in M. If $i \in N$ is matched to a dummy house, its allocation is null.
9 **return** $(x(1), \ldots, x(n))$

Lemma 2. *An allocation that maximizes welfare subject to IR has the same welfare even if IR is not imposed.*

Proof. A matching is maximum size if and only if there exist no improving path, which is an alternating path starting from an unmatched agent and ending with an unallocated house (König's theorem (see e.g., Lovász & Plummer, 2009)). Therefore if the obtained individually rational matching, M, was not maximum size then we could improve its size in such a way that all matched agents in M remain matched, contradicting with the maximality of M among IR matchings. □

As a result of the lemma, we obtain the following proposition.

Proposition 10. *MIR returns an allocation that satisfies the maximum number of agents.*

Corollary 2. *MIR returns an allocation that is Pareto optimal.*

The argument for MIR being strategyproof is exactly the same as that of MSIR being strategyproof.

Proposition 11. *MIR is strategyproof.*

For housing markets, in contrast to MSIR, MIR may not be core stable.

Proposition 12. *For 1-0 utilities, an IR welfare maximizing allocation may not be core stable.*

Proof. Consider a four agent setting in which each agent owns a house unacceptable to himself. Agent 2 owns a house that is acceptable to 4 and 1 and agent 1 owns a house that is acceptable to 2 and 3. Consider an allocation in which agent 3 gets 1's house and 4 gets 2's house. Such an allocation is IR and satisfies the maximum number of agents. However, it is not core stable because 1 and 2 can form a blocking coalition. □

In this paper, two new mechanisms called MSIR and MIR were introduced. See Table 1 for a summary of properties satisfied by the mechanisms for house allocation with existing tenants.

	MSIR	**MIR**
Strategyproof	+	+
S-IR	+	−
IR	+	+
Core stable	+	−
Pareto optimal	−	+
Max welfare subject to S-IR	+	−
Max welfare subject to IR	−	+
Max welfare	−	+

Table 1: Properties satisfied by mechanisms for house allocation with existing tenants with dichotomous preferences.

References

Abdulkadiroğlu, A., & Sönmez, T. (1999). House allocation with existing tenants. Journal of Economic Theory, 88(2), 233–60.

Abraham, D., Blum, A., & Sandholm, T. (2007). Clearing algorithms for barter exchange markets: Enabling nationwide kidney exchanges. In: Proceedings of the 8th ACM Conference on Electronic Commerce (ACM-EC), ACM Press, 295–304.

Abraham, D. J., Cechlárová, K., Manlove, D. F., & Mehlhorn, K. (2005). Pareto optimality in house allocation problems. In: Proceedings of the 16th International Symposium on Algorithms and Computation (ISAAC), Vol. 3341 of Lecture Notes in Computer Science (LNCS), 1163–75.

Aziz, H., & de Keijzer, B. (2012). Housing markets with indifferences: A tale of two mechanisms. In: Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI), 1249–55.

Biro, P., Manlove, D. F., & Rizzi, R. (2009). Maximum weight cycle packing in optimal kidney exchange programs. Discrete Mathematics, Algorithms and Applications, 1(4), 499–517.

Bogomolnaia, A., & Moulin, H. (2004). Random matching under dichotomous preferences. Econometrica, 72(1), 257–79.

Hylland, A., & Zeckhauser, R. (1979). The efficient allocation of individuals to positions. Journal of Political Econom, 87(2), 293–314.

Jaramillo, P., & Manjunath, V. (2012). The difference indifference makes in strategyproof allocation of objects. Journal of Economic Theory, 147(5), 1913–46.

Korte, B., & Vygen, J. (1999). Combinatorial optimization (5th ed.). Heidelberg, Germany: Springer-Verlag.

Krysta, P., Manlove, D. F., Rastegari, B., & Zhang, J. (2014). Size versus truthfulness in the house allocation problem. In: Proceedings of the 15th ACM Conference on Economics and Computation (ACM-EC), ACM Press, 453–70.

Lovász, L., & Plummer, M. D. (2009). Matching theory. AMS Chelsea Publishing.

Ma, J. (1994). Strategy-proofness and the strict core in a market with indivisibilities. International Journal of Game Theory, 23(1), 75–83.

Plaxton, C. G. (2013). A simple family of top trading cycles mechanisms for housing markets with indifferences. In: Proceedings of the 24th International Conference on Game Theory.

Shapley, L. S., & Scarf, H. (1974). On cores and indivisibility. Journal of Mathematical Economics, 1(1), 23–37.

Sönmez, T., & Ünver, M. U. (2010). House allocation with existing tenants: A characterization. Games and Economic Behavior, 69(2), 425–45.

Svensson, L.-G. (1994). Queue allocation of indivisible goods. Social Choice and Welfare, 11, 323–30.

Journal of Mechanism and Institution Design
ISSN: 2399-844X (Print), 2399-8458 (Online)
DOI:10.22574/jmid.2018.12.005

A REEXAMINATION OF THE COASE THEOREM

Jingang Zhao

University of Saskatchewan, Canada
j.zhao@usask.ca

ABSTRACT

This paper makes three advances: 1) It fixes the empty-core problem of the Coase theorem; 2) it provides the smallest upper bound of transaction costs below which the optimal or efficient outcomes can be achieved; and 3) it establishes two mathematical theorems that capture the main insights and major aspects of the Coase theorem. A simpler version of the theorems says that in a coalitional production economy without transaction costs, the maximal payoff will be produced by the optimal firms and be allocated in the always non-empty core.

Keywords: Coalition formation, core, optimal firms.
JEL Classification Numbers: C71, D23, L11, L23.

1. INTRODUCTION

The Coase theorem is one of the most cited, most debated and most confusing "theorems" in economics.[1] One popular version is stated as: Bargaining will lead to a Pareto efficient outcome if transaction costs are sufficiently low,

This paper replaces earlier drafts circulated under various titles, such as "The maximal payoff and coalition formation in coalitional games" (2006) and "Two mathematical versions of the Coase theorem" (2016), which had been cited in several previous studies (e.g., Arribillaga, 2010; Cesco, 2012). I thank Donald Smythe, Hao Wang, Zaifu Yang, two anonymous referees, and seminar/conference participants at Huazhong U of Sciences and Technology, Institute of System Sciences (CAS), Konan U, Kyoto U, Peking U, Renmin U of China, U of Saskatchewan, Waseda U, CEA Annual Conference, Canadian Econ Theory Conference, Coalition Theory Network Workshop, and World Congress of Game Theory Society for valuable comments and encouragements. All errors, of course, are my own.

[1] The earliest reference to (or version of) Coase theorem was in Stigler (1966, 113): "The Coase theorem thus asserts that under perfect competition private and social costs will be equal."

regardless of the initial allocation of property. As in the claim by Regan (1972, 428) that "the Coase theorem is a proposition in the theory of games, and not a proposition about traditional markets or competitive equilibrium," this paper formalizes the Coase theorem as two mathematical theorems in cooperative games and elevates it to an economic law on non-market transactions that is comparable to Smith's invisible hand on market transactions.

Precisely, this paper advances the literature on the Coase theorem in three aspects: (*i*) It fixes the purported empty-core problem of the Coase theorem (Aivazian & Callen, 1981; Coase, 1981); (*ii*) it provides the smallest upper bound of transaction costs below which the optimal or efficient outcomes can be achieved; and (*iii*) it establishes two mathematical theorems that capture the main insights and major aspects of the Coase theorem. The mathematical essence of these two theorems is the duality between sub-coalitions' producing ability (which defines the maximal payoff) and blocking power (which defines the core stability). Such duality reveals that a firm's shadow value is equal to its balancing weight, and it implies not only the three known theorems for a non-empty usual core but also three new core theorems in coalitional games.

The success in the author's search for a repair of the empty-core problem is the discovery that maximal profits in the example of Aivazian & Callen (1981), or efficient outcomes in more general cases, are achieved in advanced forms of production or minimal balanced collections of the firms rather than conventional forms of production such as the monopoly or other partitions of the firms.

The rest of the paper is organized as follows. Section 2 provides a fix for the empty-core problem in the example of Aivazian & Callen (1981). Section 3 studies the duality between sub-coalitions' blocking power and producing ability in coalitional TU (transferable utility) games. Sections 4 and 5 establish the TU and NTU (non-transferable utility) Coase theorems, respectively. Section 6 concludes, and the appendix provides proofs.

2. A FIX FOR THE PURPORTED EMPTY-CORE EXAMPLE OF THE COASE THEOREM

An important study of the Coase theorem was the following empty-core example reported in Aivazian & Callen (1981):

Example 2.1 *There are seven possible firms:*[2] $S = 1,2,3,12,13,23$, *and* 123, *and their daily profits are:* $v(1) = 3000$, $v(2) = 8000$, $v(3) = 24000$; $v(12) = 15000$, $v(13) = 31000$, $v(23) = 36000$; *and* $v(123) = 40000$.

This is an example of three-person coalitional TU game or three-owner coalitional economy. Its five partitions or conventional market structures are: $B_0 = \{1,2,3\}$, $B_1 = \{1,23\}$, $B_2 = \{2,13\}$, $B_3 = \{3,12\}$, and $B_m = \{123\}$, whose total profits satisfy: $\pi(B_0) = v(1)+v(2)+v(3) = 35000 < \pi(B_1) = v(1)+v(23) = \pi(B_2) = \pi(B_3) = 39000 < \pi(B_m) = v(123) = 40000$. Since a monopoly produces the largest profits among the five partitions, the Coase theorem implies that the monopoly or monopoly merger will be formed. However, this conclusion breaks down after one checks coalitional rationality. Let $x = (x_1,x_2,x_3) \geq 0$ be a split of $\pi(B_m) = 40000$. Rationality requires that the split has no blocking coalitions, or be in the usual core, or satisfy the following two sets of inequalities:

i) $x_1 \geq v(1) = 3000$, $x_2 \geq v(2) = 8000$, $x_3 \geq v(3) = 24000$;

ii) $x_1 + x_2 \geq v(12) = 15000$, $x_1 + x_3 \geq v(13) = 31000$, and $x_2 + x_3 \geq v(23) = 36000$.

Adding up the inequalities in (*ii*) yields $x_1 + x_2 + x_3 \geq (15000 + 31000 + 36000)/2 = 41000 > 40000$, which contradicts $x_1 + x_2 + x_3 = 40000$. Thus any split of monopoly profits is blocked by at least one coalition. Consequently, the monopoly cannot be formed so the Coase theorem fails.

Is the above argument false, or could the insight of the Coase theorem possibly be wrong? Telser (1994) observed that "Coase's elaborate analysis in his comment (1981) fails to come to grip with the issues raised by this example," and concluded that "The Coase "theorem" needs much repair when there is an empty core" (see Aivazian & Callen, 2003, for a discussion). Such a serious issue remained unsettled for more than three decades, until this study. To see the repair, move a step deeper inside this coalitional economy, assume that each of the three singleton firms is owned by a full-time worker with 8 hours of labor inputs, and that they produce a homogeneous product called profit from labor by the following linear production functions:

$f_1(x) = 3000x/8 = 375x$, $f_2(x) = 1000x$, $f_3(x) = 3000x$, $0 \leq x \leq 8$;

[2] We simplify $v(\{i\})$ as $v(i)$, $v(\{1,2\})$ as $v(12)$. Similar simplifications apply to other coalitions.

$f_{12}(x) = 937.5x$, $f_{13}(x) = 1937.5x$, $f_{23}(x) = 2250x$, $0 \leq x \leq 16$; and $f_{123}(x) = 5000x/3$, $0 \leq x \leq 24$.

The singleton, two-member and monopoly firms have 8, 16 and 24 hours of fixed labor inputs, respectively. It is easy to see that these production functions generate the same profits as those in Example 2.1. Now, consider operating each of the three two-party firms at full capacity for 4 hours (or at half capacity for 8 hours), and assume such re-organization of production does not involve any new costs (i.e., zero transaction cost). This can be arranged, for example, in the following sequence: $S = 12$ opens at full capacity from 8:00 $a.m.$-noon, $S = 13$ from noon-4:00 $p.m.$, and $S = 23$ from 4:00-8:00$p.m.$ The profits from this new or advanced form of production are truly maximal and are given by $mp = f_{12}(8) + f_{13}(8) + f_{23}(8) = [v(12) + v(13) + v(23)]/2 = 41000 > \pi(B_m) = 40000$, which exceeds the profits of operating the monopoly at full capacity for 8 hours. Define the new core as the splits of the above mp that are unblocked by all subcoalitions. One can check that the new core of Example 2.1 has a unique vector: $x_1 = 5000$, $x_2 = 10000$, $x_3 = 26000$, which is the optimal outcome predicted by the Coase theorem.[3]

Note that the above optimal outcome is achieved when all three firms open for business only half of the time and all three workers have two part-time jobs, instead of one full-time job. This conclusion, not reported in previously literature, will provide a new argument in studying labor theory and production theory.

Now, the repair is to replace conventional forms of production with advanced forms of production, or replace the usual core with the always non-empty new core (or simply, the core), so the Coase theorem is now free of the empty-core problem and will be precisely stated as Theorem 2 (with transferable utility) in section 4 and Theorem 4 (with non-transferable utility) in section 5.

[3] Related empty-core problems can be similarly fixed. Consider, for example, the empty-core in Gonzalez et al. (2016), which is caused by adding the democratic and fairness properties in their study of efficiency and neutrality (Calabresi, 1968; Regan, 1972). By replacing the grand coalition's payoff in their equation (3) with the game's maximal payoff (i.e., replacing the usual core with the new core), efficiency will always hold, and the Coase theorem will be free of the empty-core problem.

3. THE MAXIMUM OF GENERATED-PAYOFFS AND THE DUALITY IN COALITIONAL TU GAMES

Let $N = \{1, 2, \ldots, n\}$ be the set of players, $\mathbb{N} = 2^N$ be the set of all coalitions. A coalitional TU game (or a game in characteristic form) is given by

$$\Gamma = \{N, v(\cdot)\}, \tag{1}$$

which is a set function $v : \mathbb{N} \to R_+$ with $v(\varnothing) = 0$, specifying a joint payoff $v(S)$ for each coalition $S \subseteq N$. We will refer game (1) as a coalitional economy and a player i as the owner of firm i, or as a worker i, when the emphasis is on the Coase theorem. We use a lowercase v in $v(\cdot)$ to define the above TU game (1), and an uppercase V in $V(\cdot)$ to define NTU games in section 5.

Let $X(v(N) = \{x \in R^n_+ | \Sigma_{i \in N} x_i = v(N)\}$ denote the preimputation space or the set of payoff vectors that are splits of $v(N)$. A split $x = (x_1, \ldots, x_n) \in X(v(N))$ satisfies the rationality of a coalition $S \subseteq N$ or is unblocked by S if it gives S no less than $v(S)$ (i.e., $\Sigma_{i \in S} x_i \geq v(S)$), and is in the usual core[4] if it is unblocked by all $S \neq N$. Denote the usual core of (1) as

$$c_0(\Gamma) = \{x \in X(v(N)) | \Sigma_{i \in S} x_i \geq v(S) \text{ for all } S \neq N\}. \tag{2}$$

We use a lowercase c in $c_0(\Gamma)$ to denote the above TU core and an uppercase C in $C_0(\Gamma)$ to denote the NTU core in section 5.

The concept of generated-payoffs is defined by balanced collections of coalitions. Given a collection of coalitions $B = \{T_j \subseteq N | j = 1, \ldots, k\}$ and a player $i \in N$, let $B(i) = \{T \in B | i \in T\}$ denote the subset of coalitions of which i is a member. Then, B is a balanced collection if it has a balancing vector $w = \{w_T | T \in B\} \in R^k_{++}$ such that $\Sigma_{T \in B(i) W_T} = 1$ for each $i \in N$.

To see the intuition of a balancing vector, treat game (1) as a coalitional economy as in Example 2.1, where each singleton firm i has one full-time worker or 8 hours of labor inputs. Thus, each firm $S \subseteq$ N will have $8k(S)$ hours of labor inputs or $k(S)$ workers (i.e., $k(S) = |S|$ is the cardinality of S), and

[4] Note that Shapley (1955), not Gillies (1959), was the earliest known paper that defined the core. Zhao (2018) has summarized the early history of the core, from six sources, as: *i*) Gillies first used the term core during 1952-1953 referring to some intersections of the stable sets, *ii*) Shapley first formulated core solution during a conversation with Shubik in the same two-year period, and *iii*) Shapley presented his core solution in Kuhn's conference (March 1953), three months before Gillies finished his dissertation in June 1953.

produces a daily profit of $v(S)$ by operating at full capacity for 8 hours, with the following linear production function:

$$f_S(x) = v(S)x/[8k(S)], 0 \leq x \leq 8k(S). \tag{3}$$

Now, each balanced B with a balancing vector w defines the following form of production:

Operate each firm $T \in B$ at full capacity for $8w_T$ hours, and produce a profit equal to

$$gp(B) = \Sigma_{T \in B} f_T(8w_T k(T)) = \Sigma_{T \in B} w_T v(T). \tag{4}$$

The above operation and its generated payoff $gp(B)$ are feasible because the condition $\Sigma_{T \in B(i)} w_T = 1$ or balancedness ensures that each worker i works exactly 8 hours (i.e., i works for $8w_T$ hours at each T in $B(i)$). In Example 2.1, the collection $B_5 = \{12, 13, 23\}$ with balancing vector $w = (1/2, 1/2, 1/2)$ yields the payoff $gp(B_5) = [v(12) + v(13) + v(23)]/2 = 41000$.

Thus, balancing weights here are proportions of inputs that each firm $T \in B$ receives from all its owners (or that a worker i allocates to his/her firms in $B(i)$).[5] Such productions represent an advance beyond conventional forms of production or partitions, because they include partitions as special cases and they could, as in Example 2.1, possibly produce higher payoffs. The discovery of such higher payoffs beyond $v(N)$ is the key in our search for a repair for the empty-core problem of the Coase theorem.[6]

A balanced collection is minimal if no proper subcollection is balanced. It is known that a balanced collection is minimal if and only if its balancing vector is unique (Shapley, 1967). Denote the set of all minimal balanced collections (excluding the grand coalition) as

$$\Phi = \{B = \{T_1, \ldots, T_k\} | N \notin B, B \text{ is a minimal balanced collection}\}. \tag{5}$$

[5] So far we have three interpretations of the balancing weights: *i*) percentage of time during which the firm operates; *ii*) proportion of inputs that the firm receives; and *iii*) the frequency (or probability) with which the firm forms (or a player joins his/her coalition), assuming that the game is replicated/repeated for a finite number of times (or that uncertainty is added into the game). Other interpretations remain to be discovered.

[6] It is known that the grand coalition's payoff $v(N)$ might be less than the total payoffs in a non-monopoly partition or in a balanced collection, see Sun et al. (2008) on market games, Bennett (1983) on aspiration-core, and Guesnerie & Oddou (1979) on c-core. However, this study is the first to connect with the Coase theorem and treat it as a situation in which non-monopoly firms could produce more than the monopoly does.

In three-person games like Example 2.1, Φ has five entries: the four non-monopoly partitions, B_0, B_1, B_2, B_3, plus $B_5 = \{12, 13, 23\}$. We are now ready to define the maximum of generated-payoffs (mgp) and maximal payoffs (mp):

Definition 3.1 *Given game (1), $gp(B)$ and Φ in (4)-(5), mgp and mp are given by*

$$mgp = mgp(\Gamma) = Max\{gp(B)|B \in \Phi\}, \text{ and} \tag{6}$$
$$mp = mp(\Gamma) = Max\{mgp(\Gamma),\ v(N)\}. \tag{7}$$

The definition considers only minimal balanced collections because mgp is achieved among minimal balanced collections, just as the optimal value in linear programming is achieved among the extreme points.[7]

The following duality result is the theoretical foundation of our repair for the empty-core problem of the Coase theorem.

Theorem 3.1 *Given game (1), the maximization problem (6) is dual to the following minimization problem for the minimum no-blocking payoff (mnbp Zhao, 2001):*

$$mnbp = mnbp(\Gamma) = Min\{\Sigma_{i\in N} x_i | \Sigma_{i\in S} x_i \geq v(S), all\ S \neq N\}, \tag{8}$$

so $mgp(\Gamma) = mnbp(\Gamma)$ holds.

By the above duality, a firm's shadow value in (8) is equal to the balancing weight in (6), and the minimal worth of the grand coalition needed to guarantee no-blocking given in (8) is equal to the maximal payoffs produced by sub-coalitions given in (6). Because $mnbp$ represents sub-coalitions' power to block proposed splits of $v(N)$ (Zhao, 2001),[8] mgp represents their ability to produce payoffs that are different from $v(N)$, their producing ability and blocking power are dual to each other.

Note that the above linear programming problem (8) is different from earlier ones for proving the Bondareva-Shapley theorem (e.g., Bondareva 1962;

[7] It is easy to show that the above mp is equal to the grand coalition's payoff in the game's cover. When functions in (3) satisfy $f_S(x) = \Sigma_{i\in S} f_i(\lambda_i)$ for each S, all $\lambda \in R^S_+$ and $\Sigma_{i\in S}\lambda_i = x$, game (1) becomes identical to a market (T, G, A, U) as defined in Shapley & Shubik (1969), with $T = N$, $G = R^n_+$, $A = I = I_n$, and $U = \{f_i | i \in N\}$.

[8] See Watanabe & Matsubayashi (2013) and Karsten & Basten (2014) for applications of $mnbp$.

Shapley 1967; Myerson 1991, 432) in that the grand coalition's constraint is removed in (8) (i.e., all $S \neq N$) while it was included in earlier ones (i.e., all $S \subseteq N$). Needless to say, it is the removal of the grand coalition's constraint in Zhao (2001) and in this study that distinguishes the author's approach from earlier studies, which leads to the discovery of the duality between the producing ability and the blocking power of all sub-coalitions. Such a duality is perhaps the most salient property in cooperative games, because it has three important advantages as discussed below.
First, it also holds in NTU games (see Theorem 3 in section 5). Second, it implies three existence theorems on the usual core (two known and one new), as given in the next corollary.

Corollary 3.2 *Given game (1), the following four statements are equivalent: i) its usual core is non-empty; ii) the game is balanced (Bondareva, 1962; Shapley, 1967); iii) the grand coalition's payoff is large enough to guarantee no-blocking (Zhao, 2001); and iv) players cannot produce a higher payoff than the grand coalition's payoff.*

Precisely, the above core arguments in parts *ii-iv*) are: *ii*) $\Sigma_{T \in B} w_T v(T) \leq v(N)$ for each balanced B with a balancing vector w, *iii*) $v(N) \geq mnbp(\Gamma)$, and *iv*) $mgp(\Gamma) \leq v(N)$.

Third, Theorem 1 also implies, as shown in the next section, the answers to four important questions arising from the Coase theorem: What payoffs will be split? How will the payoff be split? What firms will form? and How much inputs will each of the formed firms receive? Note that the Coase theorem won't be complete if any of these four questions is not answered.

4. THE TU COASE THEOREM

By the new core argument in part *iv*) of Corollary 1, players in games with an empty usual core will not split the grand coalition's payoff $v(N)$, which is smaller than the game's mgp. Then, what payoffs will they split? We postulate that they split the maximal payoff $mp(\Gamma)$ in (7). By $mp(\Gamma) = v(N \geq mgp(\Gamma)$ if $c_0(\Gamma) \neq \varnothing$, and $mp(\Gamma) = mgp(\Gamma) > v(N)$ if $c_0(\Gamma) = \varnothing$, it stands to reason that they will always split $mp(\Gamma)$. This answers the question of what payoffs will be split.

Next, consider the question of how to split the maximal payoff. Let the set of minimal solutions for $mnbp(\Gamma)$ in (8) be denoted by

$$Y = Y(\Gamma) = Arg - Min\{\Sigma_{i\in N} x_i | \Sigma_{i\in S} x_i \geq v(S), \text{ all } S \neq N\}. \tag{9}$$

Coalitional rationality leads to the new core or simply the core given by $c(\Gamma) =$

$$\{x \in X(mp) | \Sigma_{i\in S} x_i \geq v(S), \text{all } S \subseteq N\} = \begin{cases} c_0(\Gamma) \text{ if } v(N) = mp(\Gamma) \\ Y(\Gamma) \text{ if } v(N) < mp(\Gamma), \end{cases} \tag{10}$$

which answers the question of how to split $mp(\Gamma)$. Note that the above core is always nonempty because it is identical to the usual core in (2) when the usual core is nonempty, and the set of minimal solutions in (8) or (9) when the usual core is empty.

It is useful to note that the above core in (10) includes the usual core of the game's cover (or balanced cover, Shapley, 1967), which is a new game $\Gamma_{bc} = \{N, \overline{v}(\cdot)\}$, where each $\overline{v}(S) = mp(\Gamma_S)$ is the maximal payoff of the subgame $\Gamma_S = \{S, v(\cdot)\}$.[9]

Now, consider the question of what firms will form. Because players always split mp, they will form the optimal collections that generate mp. Let the set of optimal solutions for $mgp(\Gamma)$in (6) be given by

$$B_0 = B_0(\Gamma) = \{B \in \Phi | gp(B) = mgp(\Gamma)\} = Arg - Max\{gp(B) | B \in \Phi\}. \tag{11}$$

Then, the set of optimal collections can be given as

$$B^* = B^*(\Gamma) = \begin{cases} \{N\} \text{ if } mgp(\Gamma) < v(N) \\ B_0(\Gamma) \text{ if } mgp(\Gamma) > v(N) \\ \{N\} \cup B_0(\Gamma) \text{ if } mgp(\Gamma) = v(N), \end{cases} \tag{12}$$

which answers the question of what firms will form.

Finally, the unique balancing vector for each optimal B answers the question of how much inputs will each of the formed firms receive. The above answers cover all the relevant aspects of the Coase theorem for our coalitional economy (1), which can now be stated as:

[9] The author thanks the referee for pointing out that the core in (10) is identical to the aspiration-core (Bennett, 1983). Because the duality in Theorem 1 yields a new understanding about the internal structure of the core in (10), and because the aspiration-core does not extend to NTU games and is not related to the Coase theorem, the term "core" or "new core" is kept for the set of solutions given in (10).

Theorem 4.1 *In coalitional economy* (1) *without transaction costs, owners will produce the maximal payoff by allocating inputs to optimal firms based on the firms' shadow values, and they will split the maximal payoff within the non-empty core.*

Precisely, the maximal payoff, core and optimal firms are respectively given in (7), (10), and (12), and the balancing vector for each set of optimal firms in (12) specifies their shadow values or proportions of inputs received from the owners. In Example 2.1, one has $mp(\Gamma) = mgp(\Gamma) = 41000 > v(N) = 40000$. By Theorem 2, the optimal collection or the set of optimal firms with input allocations (or operating hours) is $B^*(\Gamma) = B_0(\Gamma) = \{12, 13, 23\}$ with $w = (1/2, 1/2, 1/2)$, and the core is $c(\Gamma) = Y(\Gamma) = (5000, 10000, 26000)\}$.

One advantage of Theorem 2 is that it provides a method to estimate the bounds of transaction costs (or merging costs) for each firm or coalition S, using the approach introduced by the author (Zhao, 2009). For simplicity, let the merging costs for monopoly be $\tau_N > 0$ and that for all sub-coalitions be zero in our game (1). Because the owners now can only split $v(N)$-τ_N) and $mgp = mnbp$ remains unchanged, monopoly formation now requires:

$$v(N) - \tau_N \geq mnbp. \tag{13}$$

Thus, $\tau_N \leq (v(N) - mnbp)$ ($\tau_N > [v(N) - mnbp]$) holds for a successful (failed) monopoly formation.

In other words, the difference between $v(N)$ and $mnbp$ serves as an upper (lower) bound of monopoly's transaction costs below (above) which monopoly formation is possible (impossible), which can be empirically estimated.

5. THE NTU COASE THEOREM

This section studies the efficient payoffs and efficient firms in an NTU coalitional economy. An NTU coalitional economy (also called NTU bargaining economy, NTU coalitional game, and NTU game in characteristic form) is defined as

$$\Gamma = \{N, V(\cdot)\}, \tag{14}$$

which specifies, for each $S \subseteq N$, a non-empty set of payoffs $V(S)$ in $\mathbf{R}^S$, which is the Euclidean space whose dimension and coordinates are the number of

players in S and their payoffs. Let the (weakly) efficient set of each $V(S)$ be denoted by

$$\partial V(S) = \{\, y \in V(S) |\text{ there is no } x \in V(S) \text{ such that } x >> y\},$$

where vector inequalities are defined by: $x \geq y \Leftrightarrow x_i \geq y_i$, all i; $x > y \Leftrightarrow x \geq y$ and $x \neq y$; and $x >> y \Leftrightarrow x_i > y_i$, all i.

Scarf (1967b) introduced the following two assumptions for (14): *i*) each $V(S)$ is closed and comprehensive (i.e., $y \in V(S), u \in R^S$ and $u \leq y$ imply $u \in V(S)$); *ii*) for each S, $\{y \in V(S) | y_i \geq \partial V(i),\text{ all } i \in S\}$ is non-empty and bounded, where $\partial V(i) = \max\{x_i | x_i \in V(i)\} > 0$. Under these two assumptions, each $\partial V(S)$ is closed, non-empty and bounded.

Given $S \subseteq N$, a payoff vector $u \in R^n$ is blocked by S if S can obtain a higher payoff for each of its members than that given by u, or precisely if there is $y \in V(S)$ such that $y_S >> u_S = \{u_i | i \in S\}$ or $u_S \in V(S) \backslash \partial V(S)$. A payoff vector $u \in \partial V(N)$ is in the usual core if it is unblocked by all $S \neq N$, so the usual core of (14) can be given as

$$C_0(\Gamma) = \{u \in \partial V(N) | u_S \notin V(S) \backslash \partial V(S),\text{ all } S \neq N\}. \tag{15}$$

Balanced NTU games can be defined geometrically as below. For each $S \neq N$, let $\widetilde{V}(S) = V(S) \times R^{-S} \subset R^n$, where $R^{-S} = \Pi_{i \notin S} R^i$. For each minimal balanced collection $B \in \Phi$ in (5), let

$$GP(B) = \cap_{S \in B} \widetilde{V}(S), \text{ and } GP(\Gamma) = \cup_{B \in \Phi} GP(B) \tag{16}$$

denote the payoffs generated by B and the set of generated-payoffs. Note that $GP(B)$ is simplified to $GP(B) = \Pi_{S \in B} V(S)$ when B is a partition. Similar to the TU case, we only need to consider minimal balanced collections because non-minimal balanced collections don't generate additional payoffs. Now, we are ready to define the efficient generated-payoffs $EGP = EGP(\Gamma) = \partial GP(\Gamma)$, and the efficient payoffs $EP = EP(\Gamma) = \partial(GP(\Gamma) \cup V(N))$, which are the NTU counterparts of mgp and mp in (6-7).

Definition 5.1 *Given game* (14) *and its GP in* (16)*, its EGP and EP are given by*

$$EGP = \{y \in GP(\Gamma) | \exists \textit{ no } x \in GP(\Gamma) \textit{ such that } x >> y\}, \textit{ and} \tag{17}$$

$$EP = \{y \in GP(\Gamma) \cup V(N) | \exists \textit{ no } x \in GP(\Gamma) \cup V(N) \textit{ with } x >> y\}. \tag{18}$$

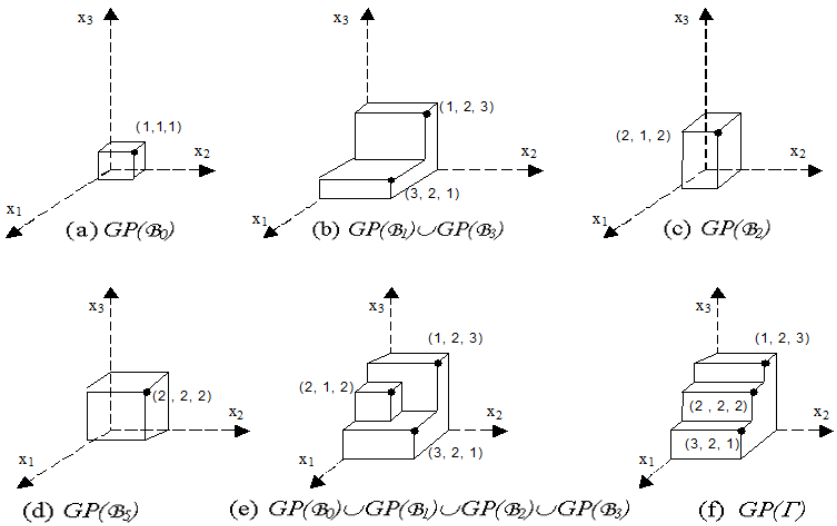

Figure 1: The generated payoffs in Example 5.1, where $B_0 = \{1,2,3\}$, $B_1 = \{1,23\}$, $B_2 = \{2,13\}$, $B_3 = \{3,12\}$ and $B_5 = \{12,13,23\}$.

Readers are encouraged to visualize the generated-payoffs in the following example, which are illustrated in Figure 1.

Example 5.1 *$n = 3$, $V(i) = \{x_i | x_i \leq 1\}$, $i = 1,2,3$; $V(12) = \{(x_1,x_2)|(x_1,x_2) \leq (3,2)\}$, $V(13) = \{(x_1,x_3)|(x_1,x_3) \leq (2,2)\}$, $V(23) = \{(x_2,x_3)|(x_2,x_3) \leq (2,3)\}$, $V(123) = \{x|x_1 + x_2 + x_3 \leq 5\}$. Let B_i, $i = 0,1,2,3$, be as in Example 2.1, and $B_5 = \{12,13,23\}$. Then, $GP\ (B_0) = \{x|x \leq (1,1,1)\}$, $GPB_1) = \{x|x \leq (1,2,3)\}$, $GPB_2) = \{x|x \leq (2,1,2)\}$, $GPB_3) = \{x|x \leq (3,2,1)\}$, and $GPB_5) = \{x|x \leq (2,2,2)\}$.*

Now, the game (14) is balanced if $GP(\Gamma) \subset V(N)$ or if for each balanced B, $u \in V(N)$ must hold if $u_S \in V(S)$ for all $S \in B$. To see a balanced game geometrically, visualize that one is flying over a city. Treat the generated-payoffs $GP(\Gamma)$ as trees and buildings in the city and the grand coalition's payoff $V(N)$ as clouds. Then, a game is balanced if one sees only clouds and unbalanced if one sees at least one building or tree top above the clouds. In Figure 2b for Example 5.1, one sees three building tops above the clouds so the game is unbalanced; in Figure 2a for the following Example 5.2, one sees only clouds so the game is now balanced.

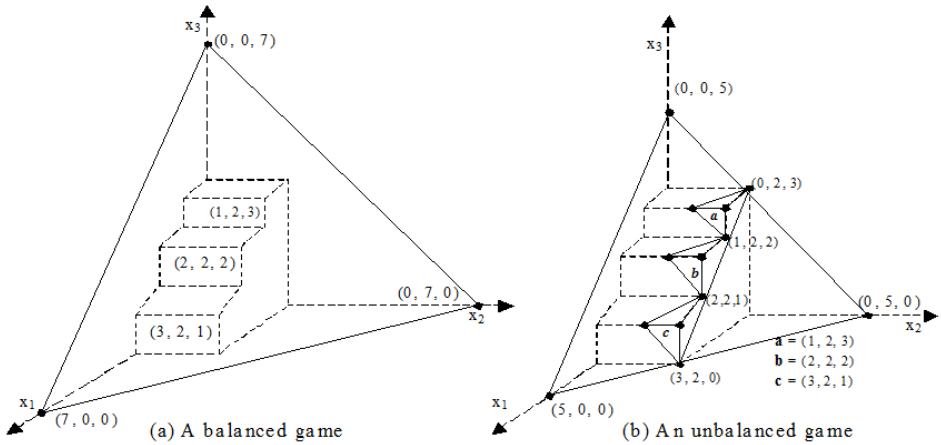

Figure 2: Balanced and unbalanced games.

Example 5.2 *Same as Example 5.1, except* $V(123) = \{x | x_1 + x_2 + x_3 \leq 7\}$.

Note that the collection $B_5 = \{12, 13, 23\}$ in Example 5.1 generates new payoffs that are outside of those generated by the four partitions and are better than $V(N)$ as shown by the difference between (e) and (f) in Figure 1. Needless to say, it is the discovery of such new and better payoffs that gives rise to our mathematical versions of the Coase theorem.

Recall that a payoff vector u is unblocked by S if $u \in [V(S) \backslash \partial V(S)]^C \times R^{-S}$ or if $u_S \notin V(S) \backslash \partial V(S)$, where the superscript C denotes the complement of a set. The following concept of minimum no-blocking frontier is the NTU counterpart of *mnbp* in (8):

Definition 5.2 *Given game* (14)*, the set of payoffs unblocked by all* $S \neq N$ $(UBP = UBP(\Gamma))$ *and the minimum no-blocking frontier* $(MNBF = MNBF(\Gamma) = \partial UBP)$ *are respectively given by*

$$UBP = \cap_{S \neq N} \{[V(S) \backslash \partial V(S)]^C \times R^{-S}\} \subset R^n, \text{ and} \tag{19}$$

$$MNBF = \{y \in UBP(\Gamma) | \exists \text{ no } x \in UBP(\Gamma) \text{ with } x << y\}. \tag{20}$$

It is easy to see that each payoff vector on or above $MNBF(\Gamma)$ is unblocked by all $S \neq N$, and the usual core can be given as $C_0(\Gamma) = UBP(\Gamma) \cap \partial V(N) = MNBF(\Gamma) \cap \partial V(N)$. Similar to the TU case, $MNBF$ represents sub-coalitions'

power to block the grand coalition's proposals. Theorem 3 below shows that sub-coalitions' blocking power and producing ability are also dual to each other in coalitional NTU games, which is the NTU counterpart of Theorem 1.

Theorem 5.1 *Given game* (14)*, its minimum no-blocking frontier and efficient generated-payoffs have a non-empty intersection.*

To put it differently, the NTU counterpart of $mnbp = mgp$ in game (1) is that the set of unblocked and efficient generated-payoffs is non-empty, or precisely,

$$Z(\Gamma) = MNBF(\Gamma) \cap EGP(\Gamma) \neq \varnothing. \tag{21}$$

It is straightforward to verify $a, b, c \in Z(\Gamma)$ in Example 5.1 (see Figure 2b), where $a = (1,2,3)$, $b = (2,2,2)$, and $c = (3,2,1)$, so $Z(\Gamma) \neq \varnothing$ holds in the example.

Recall that $EGP(\Gamma) \subseteq V(N)$ holds in balanced games. Then, $MNBF(\Gamma) \cap EGP(\Gamma) \neq \varnothing$ implies $MNBF(\Gamma) \cap \partial V(N) = C_0(\Gamma) \neq \varnothing$in balanced games. Hence, the above duality implies Scarf's core theorem Scarf (1967b). It also implies two new existence theorems on the usual core, which are summarized in the following corollary:

Corollary 5.2 *Given game* (14)*, the following three claims hold: i) its usual core is non-empty if it is balanced (Scarf, 1967b); ii) its usual core is non-empty if and only if the grand coalition's payoff is large enough to guarantee no blocking; and iii) its usual core is non-empty if players can't produce better payoffs than the grand coalition's payoff.*

The above results in parts *i*) and *iii*) can be precisely stated in one argument: $C_0(\Gamma) \neq \varnothing$ if $GP(\Gamma) \subset V(N)$, and the result in part *ii*) can be precisely stated as: $C_0(\Gamma) \neq \varnothing \Leftrightarrow$ there exists $x \in \partial V(N)$and $y \in MNBF(\Gamma)$ such that $x \geq y$.

Due to the generality of non-transferable utilities, the above NTU core results are more general than the earlier TU results in at least three aspects: *i*) the usual NTU core is no longer convex, as shown in Figure 3 for Examples 5.1-5.2; *ii*) balancedness is only sufficient and no longer necessary for a non-empty usual NTU core, see Figure 3a for a non-empty usual NTU core in the unbalanced Example 5.1; and *iii*) "players can't produce better payoffs than $V(N)$" is no longer necessary for a non-empty usual NTU core.

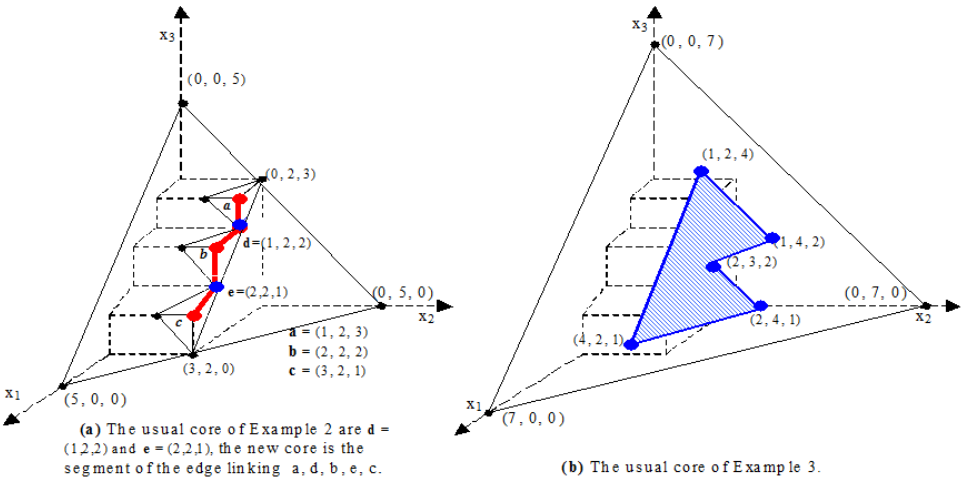

Figure 3: The usual core and the new core: blue-colored payoffs are the usual core, and red-colored payoffs are the new core with the usual core removed.

Recall that $Z(\Gamma)$ in (21) is the set of unblocked and efficient generated-payoffs. Let the subset of $Z(\Gamma)$ that are also unblocked by the grand coalition N be denoted by

$$Z^*(\Gamma) = Z(\Gamma) \cap [V(N) \backslash \partial V(N)]^C. \tag{22}$$

Now, let the set of minimal balanced collections that generate $Z(\Gamma)$ and $Z^*(\Gamma)$ be denoted respectively by

$$D_0(\Gamma) = \{B \in \Phi | GP(B) \in Z(\Gamma)\}, \text{ and} \tag{23}$$

$$D_1(\Gamma) = \{B \in D_0(\Gamma) | GP(B) \in Z^*(\Gamma)\}. \tag{24}$$

Note that $D_0(\Gamma)$ is the NTU counterpart of the minimal set $B_0(\Gamma)$ in (11). Now, by replacing the grand coalition's payoff $V(N)$ with the efficient payoffs $EP(\Gamma)$ in (18), one gets the following new NTU core (or simply the NTU core) $C(\Gamma)$ and the set of efficient firms $D^*(\Gamma)$:

$$C(\Gamma) = \{u \in EP(\Gamma) | u_S \notin V(S) \backslash \partial V(S), \text{ all } S \subseteq N\} \tag{25}$$

$$= \begin{cases} C_0 \text{ if } GP \subseteq V(N) \\ Z \text{ if } V(N) \subseteq GP \backslash \partial GP \text{ or if } V(N) \subsetneq GP \backslash \partial GP, GP \subsetneq V(N), C_0 = \varnothing \\ C_0 \cup Z^* \text{ if } V(N) \subsetneq GP \backslash \partial GP, GP \subsetneq V(N), C_0 \neq \varnothing, \end{cases}$$

$$D^*(\Gamma) = \text{The set of efficient firms or set of efficient collections} \tag{26}$$
$$= \begin{cases} \{N\} \text{ if } GP \subseteq V(N) \\ D_0 \text{ if } V(N) \subseteq GP \backslash \partial GP \text{ or if } V(N) \subsetneq GP \backslash \partial GP, GP \subsetneq V(N), C_0 = \varnothing \\ \{N\} \cup D_1 \text{ if } V(N) \subsetneq GP \backslash \partial GP, GP \subsetneq V(N), C_0 \neq \varnothing, \end{cases}$$

where $C_0 = C_0(\Gamma)$, $GP = GP(\Gamma)$, $Z = Z(\Gamma)$, $Z^* = Z^*(\Gamma)$, $D_0 = D_0(\Gamma)$ and $D_I = D_I(\Gamma)$ are, respectively, the usual core in (15), generated payoffs in (16), unblocked and efficient generated-payoffs in (21), efficient generated-payoffs that are also unblocked by grand coalition in (22), collections supporting $Z(\Gamma)$in (23), and collections supporting $Z^*(\Gamma)$ in (24). In words, the NTU core is characterized in three cases and it is equal to:

i) the usual NTU core if the game is balanced;

ii) the set of unblocked and efficient generated-payoffs if players can produce better payoffs than $V(N)$ or if players cannot produce better payoffs than $V(N)$ and the game is unbalanced with an empty usual core; and

iii) the union of the usual NTU core and a subset of unblocked and efficient generated-payoffs (i.e., $C_0 \cup Z^*$) if players cannot produce better payoffs than $V(N)$ and the game is unbalanced with a non-empty usual NTU core.

Now, our NTU Coase theorem, comprising the above answers, can be stated as:

Theorem 5.3 *In coalitional NTU economy* (14) *without transaction costs, owners will produce the efficient payoffs by allocating inputs to efficient firms based on the firms' shadow values, and will choose an efficient payoff vector from the non-empty NTU core.*

Precisely, the efficient payoffs, core payoffs and efficient firms are given in (18), (25) and (26), respectively. Figure 3a illustrates the difference between the usual NTU core (i.e., points d and e) and new NTU core (i.e., the segment of the edge linking all three peaks) in Example 5.1. Note that efficient firms in (26) are defined according to the three cases of the core in (25). There are five sets of efficient firms in Example 5.1: $B_1 = \{23,1\}$, $B_2 = \{13,2\}$, $B_3 = \{12,3\}$, $B_m = \{N\}$ and $B_5 = \{12,13,23\}$. Keep in mind that efficiency here is defined by weakly efficient solutions. For example, the payoff $(2,1,2)$ is only weakly efficient; it is not efficient in the sense of Pareto because it is Pareto-dominated by $(2,2,2)$.

Analogous to the TU case, the strong conclusion of NTU Coase theorem results from the advantages of utilizing generated-payoffs: in the usual NTU

core, players just choose from $\partial V(N)$; whereas in the new NTU core, players choose from the game's efficient payoffs, which are sometimes better than $\partial V(N)$.

6. CONCLUSION AND DISCUSSION

The above analysis has explored the possibility that owners in a coalitional economy sometimes could produce better payoffs than the monopoly payoff. It has revealed that a firm's shadow value is equal to its balancing weight, and it has captured the major aspects and main insights of the Coase theorem by establishing two mathematical theorems in coalitional economies.

By modeling non-market allocation of resources as a coalitional economy or bargaining economy, the paper not only has advanced coalition formation from partitions to minimal balanced collections but also has advanced the study of the Coase theorem in three areas. First, our two versions of the Coase theorem (i.e., Theorems 2 and 4) show that it is sometimes socially optimal for firms to shut down parts of their operations and for workers to have two or more part-time jobs. This conclusion, previously unreported, will provide a new line of argument in studying labor theory, production theory, and other related fields in economics.

Second, our two versions show precisely how the size of transaction costs in each merger or coalition could prevent or allow its formation, and this provides a two-step procedure for empirically estimating the size of transaction costs involved in each previous or future application of the Coase theorem: i) identify the merger (or the parties in the transaction problem under investigation) and convert it into a coalitional economy, and ii) compute its minimum no-blocking payoff ($mnbp$). The difference between the merger's payoff and its $mnbp$ is the estimated upper (lower) bound of transaction costs below (above) which the optimal outcome predicted by the Coase theorem holds (fails).

Finally, our two versions have the potential to open doors for applying the Coase theorem not only to TU and NTU transaction problems but also to all non-transaction problems that are modeled by coalitional games.

Appendix

Proof of Theorem 1: For each $S \neq N$, let $e_S = (x_1, \ldots, x_n)' \in \mathbf{R}^n_+$ be its incidence vector or the column vector such that $x_i = 1$ if $i \in S$ and $x_i = 0$ if $i \notin S$, and $e = e_N = (1, \ldots, 1)'$ be a column vector of ones. Then, the dual problem for the minimization problem (8) is the following maximization problem:

$$Max\{\Sigma_{S \neq N} y_S v(S) | y_S \geq 0, \text{ all } S \neq N; \text{ and } \Sigma_{S \neq N} y_S e_S \leq e\}. \tag{27}$$

We will show that (27) is equivalent to the maximization problem (6). First, we show that the inequality constraints in (27) can be replaced by equation constraints.

Let $Ay \leq e$ and $y \geq 0$ denote the constraints in (27), where $A = A_{n \times (2^n - 2)} = [e_S | S \neq N]$ is the constraint matrix, and y is the $(2^n - 2)$ dimensional vector whose indices are the proper coalitions. Let the rows of A be $a_1, \ldots, a_n$, and for each feasible y, let $T = T(y) = \{i | a_i y < 1\}$ be the set of loose constraints, so $N \backslash T = \{i | a_i y = 1\}$ is the set of binding constraints.

If $T(y) \neq \varnothing$, let z be defined as: $z_S = y_S + (1 - a_i y)$ if $S = \{i\}$, for each $i \in T$, and $z_S = y_S$ if $S \neq \{i\}$ for all $i \in T$. One sees that $z > y$ and $T(z) = \varnothing$. Hence, for any y with $T(y) \neq \varnothing$, there exists $z \geq 0, Az = e$ such that $\Sigma_{S \neq N} y_S v(S) \leq \Sigma_{S \neq N} z_S v(S)$. Thus, the feasible set of (27) can be reduced to $\{z | z \geq 0, Az = e\}$, without affecting the maximum value. So the maximization problem in (27) is equivalent to the following problem:

$$Max\{\Sigma_{S \neq N} y_S v(S) | Ay = e \text{ and } y \geq 0\}. \tag{28}$$

Next, we establish the one-to-one relationship between the extreme points of (28) and the minimal balanced collections. Note that for each feasible y in (28), $B(y) = \{S | y_S > 0\}$ is a balanced collection. Let y be an extreme point of (28). We now show that $B(y) = \{S | y_S > 0\}$ is a minimal balanced collection.

Assume by way of contradiction that $B(y)$ is not minimal, then there exists a balanced subcollection $B \subset B(y)$ with balancing vector z. Note that $z_S > 0$ implies $y_S > 0$. Therefore, for a sufficiently small $t > 0$ (e.g., $0 < t \leq 1/2$, and $t \leq Min\{y_S / |z_S - y_S| |$ all S with $y_S \neq z_S\}$), one has

$$w = y - t(y - z) \geq 0, w' = y + t(y - z) \geq 0.$$

$Ay = e$ and $Az = e$ lead to $Aw = e$ and $Aw' = e$. But $y = (w + w')/2$ and $w \neq w'$ contradict the assumption that y is an extreme point. So $B(y)$ must be minimal.

Now, let $B = \{T_1, \ldots, T_k\}$ be a minimal balanced collection with a balancing vector z. We need to show that z is an extreme point of (28). Assume again by way of contradiction that z is not an extreme point, so there exist feasible $w \neq w\prime$ in (28) such that $z = (w + w')/2$. By $w \geq 0$, $w' \geq 0$, one has

$$\{S|w_S > 0\} \subseteq B = \{S|z_S > 0\}, \text{and } \{S|w'_S > 0\} \subseteq B = \{S|z_S > 0\}.$$

The above two expressions show that both w and w' are balancing vectors for some subcollections of B. Because B is minimal, one has $w = w' = z$, which contradicts $w \neq w'$. Therefore, z must be an extreme point of (28).

Finally, by the standard results in linear programming, the maximal value of (28) is achieved among the set of its extreme points, which are equivalent to the set of the minimal balanced collections, so (28) is equivalent to $Max\{\Sigma_{S \in B} y_S v(S) | B \in \Phi\}$, where $B \in \Phi$ in (5) is a minimal balanced collection with the balancing vector y. This shows that (27) is equivalent to the maximization problem (6) for mgp, which completes the proof. **Q.E.D**

Proof of Corollary 1: The equivalences follow from Theorem 1. Note that part *ii*) was established in previous studies using the duality theorem for $Min\{\Sigma_{i \in N} x_i | \Sigma_{i \in S} x_i \geq v(S), all\ S \subseteq N\}$. **Q.E.D**

Proof of Theorem 2: Discussions before the theorem serve as a proof. **Q.E.D**

Our proof for Theorem 3 uses the following lemma on open covering of the simplex $\Delta^N = X(1) = \{x \in \mathbf{R}^n_+ | \Sigma_{i \in N} x_i = 1\}$.

Lemma 1 (Scarf, 1967a; Zhou, 1994): *Let $\{C_S\}$, $S \neq N$, be a family of open subsets of Δ^N that satisfy $\Delta_{N \setminus \{i\}} = \{x \in \Delta^N | x_i = 0\} \subset C_{\{i\}}$ for all $i \in N$, and $\cup_{S \neq N} C_S = \Delta^N$, then there exists a balanced collection of coalitions B such that $\cap_{S \in B} C_S \neq \varnothing$.*

Proof of Theorem 3: Let UBP be the set of unblocked payoffs in (19), and EGP be the boundary or (weakly) efficient set of the generated payoff in (17). We shall first show that $UBP \cap EGP \neq \varnothing$.

For each coalition $S \neq N$, let $W_S = \{IntV(S) \times R^{-S}\} \cap EGP$ be an open (relatively in EGP) subset of EGP, where $IntV(S) = V(S) \setminus \partial V(S)$ is the interior of $V(S)$. For each minimal balanced collection of coalitions B, we claim that

$$\cap_{S \in B} W_S = \varnothing \tag{29}$$

holds. If (29) is false, there exists $y \in EGP$ and $y \in IntV(S) \times \mathbf{R}^{-S}$ for each $S \in B$. We can now find a small $t > 0$ such that $y + te \in IntV(S) \times \mathbf{R}^{-S}$ for

each $S \in B$, where e is the vector of ones. By the definitions in (16)-(17), $y+te \in GP(B) = \cap_{S\in B}\{V(S) \times \mathbf{R}^{-S}\} \subset GP$, which contradicts $y \in EGP$. This proves (29).

Now, suppose by way of contradiction that $UBP \cap EGP = \varnothing$. Then, $EGP \subset UBP^C$, where superscript C denotes the complement of a set. The definition of W_S and

$$UBP^C = \{\cap_{S\neq N}\{[V(S)\backslash\partial V(S)]^C \times \mathbf{R}^{-S}\}\}^C = \cup_{S\neq N}\{IntV(S) \times \mathbf{R}^{-S}\}$$

together lead to $\cup_{S\neq N}W_S = EGP$, so $\{W_S\}, S \neq N$, is an open cover of EGP.

Because the set of generated payoffs is comprehensive and bounded from above, and the origin is in its interior (by $\partial V(i) > 0$, all i), the following mapping from EGP to Δ^N:

$$f : x \to x/\Sigma x_i,$$

is a homeomorphism. Define $C_S = f(W_S)$ for all $S \subseteq N$, one sees that $\{C_S\}, S \neq N$, is an open cover of $\Delta^N = f(EGP)$.

For each $i \in N$, $\partial V(i) > 0$ leads to $EGP \cap \{x \in \mathbf{R}^n | x_i = 0\} \subset W_{\{i\}}$, which in turn leads to $\Delta_{N\backslash\{i\}} = \{x \in \Delta^N | x_i = 0\} = f(EGP \cap \{x \in \mathbf{R}^n | x_i = 0\}) \subset C_{\{i\}} = f(W_{\{i\}})$. Therefore, $\{C_S\}$, $S \neq N$, is an open cover of Δ^N satisfying the conditions of Lemma 1, so there exists a balanced collection of coalitions B_0 such that $\cap_{S\in B_0}C_S \neq \varnothing$, which leads to $\cap_{S\in B_0}W_S \neq \varnothing$. This contradicts (29). Hence, $UBP \cap EGP \neq \varnothing$.

For each $x \in UBP \cap EGP$, we claim $x \in MNBF$. If this is false, we can find a small $\tau > 0$ such that $x - \tau e \in UBP$. Let $B \in \Phi$ be the minimal balanced collection of coalitions such that $x \in GP(B) = \cap_{S\in B}\{V(S) \times \mathbf{R}^{-S}\}$. Then, $x - \tau e \in IntV(S) \times \mathbf{R}^{-S}$ for each $S \in B$, which contradicts $x - \tau e \in UBP$. Therefore, $MNBF \cap EGP = UBP \cap EGP \neq \varnothing$. **Q.E.D**

Proofs of Corollary 2 and Theorem 4: The discussions preceding the corollary and the theorem serve as their proofs. **Q.E.D**

References

Aivazian, V., & Callen, J. (1981). The Coase theorem and the empty core. Journal of Law and Economics, 24, 175–81.

Aivazian, V., & Callen, J. (2003). The core, transaction costs and the coase theorem. Constitutional Political Economy, 14, 287–99.

Arribillaga, R. (2010). The balanced core in NTU games: Axiomatic characterization. Instituto de Matemática Aplicada San Luis (UNSL-CONICET), Working Paper.

Bennett, E. (1983). The aspiration approach to predicting coalition formation and payoff distribution in side payment games. International Journal of Game Theory, 12, 1–28.

Bondareva, O. (1962). The theory of the core in an n-person game (in Russian). Vestnik Leningrad. Univ, 13, 141–2.

Calabresi, G. (1968). Transaction costs, resource allocation and liability rules: A comment. Journal of Law and Economics, 11, 67–73.

Cesco, J. (2012). Nonempty core-type solutions over balanced coalitions in TU-games. International Game Theory Review, 14, 1250018.1-16.

Coase, R. H. (1960). The problem of social cost. Journal of Law and Economics, 3, 1–44.

Coase, R. H. (1981). The Coase theorem and the empty core: A comment. Journal of Law and Economics, 24, 183–87.

Gillies, D. (1959). Solutions to general non-zero-sum games. In A. Tucker & R. Luce (Eds.), Contributions to the theory of games IV. Princeton, NJ: Princeton University Press.

Gonzalez, S., Marciano, A., & Solal, P. (2016). The social cost problem, rights and the (non) empty core. Working Paper, Groupe d'Analyse et de Théorie Economique, UJM Saint-Etienne.

Guesnerie, R., & Oddou, C. (1979). On economic games which are not necessarily superadditive. Economics Letters, 3, 301–6.

Karsten, F., & Basten, R. (2014). Spare parts inventory pooling: How to share the benefits? European Journal of Operational Research, 233, 94–104.

Myerson, R. (1991). Game theory. Cambridge, MA: Harvard University Press.

Regan, D. (1972). The problem of social cost revisited. Journal of Law and Economics, 15, 427–37.

Scarf, H. (1967a). The approximation of fixed points of a continuous mapping. SIAM Journal of Applied Mathematics, 15, 1328–42.

Scarf, H. (1967b). The core of an n-person game. Econometrica, 35, 50–69.

Shapley, L. (1955). Markets as cooperative games. Rand Corporation Papers, P-629, 1–5.

Shapley, L. (1967). On balanced sets and cores. Naval Research Logistics Quarterly, 14, 453–60.

Shapley, L., & Shubik, M. (1969). On market games. Journal of Economic Theory, 1, 9–25.

Stigler, G. (1966). Theory of price (3rd ed.). New York, NY: Macmillan.

Sun, N., Trockel, W., & Yang, Z. (2008). Competitive outcomes and endogenous

coalition formation in an n-person game. Journal of Mathematical Economics, 44, 853–60.

Telser, L. (1994). The usefulness of core theory in economics. Journal of Economic Perspectives, 8, 151–64.

Watanabe, T., & Matsubayashi, N. (2013). Note on stable mergers in a market with asymmetric substitutability. Economics Bulletin, 33, 2024–33.

Zhao, J. (2001). The relative interior of base polyhedron and the core. Economic Theory, 18, 635–48.

Zhao, J. (2009). Estimating merging costs by merger preconditions. Theory and Decision, 66, 373–99.

Zhao, J. (2018). Three little-known and yet still significant contributions of Lloyd Shapley. Games and Economic Behavior, 108, 592–99.

Zhou, L. (1994). A new bargaining set of an n-person game and endogenous coalition formation. Games and Economic Behavior, 6, 512–26.

How to prepare a paper for submission

Before submitting a paper to **this Journal**, the authors are advised to follow closely the following instructions to prepare the paper.

1. Papers submitted to this Journal must be unpublished and original work that is neither under review elsewhere nor will be submitted elsewhere for publication without withdrawing from this Journal.

2. This Journal requires that all results (experimental, empirical and computational) be replicable. All underlying data necessary to replicate results must be made available to the Journal.

3. Papers must be submitted in electronic form (preferably as pdf files) and the size of the text font must be at least 12 point. Each figure and table must be included on the relevant page of the paper and should not be collected at the end of the paper. The list of references should appear after any appendices, as the last part of the paper.

4. There is no restriction on the number of pages of any submitted manuscript. We seek to process any first submission within 3 months. However, very long manuscripts may take considerably more time to review.

5. Each submitted paper should have an abstract of no more than 150 words containing no mathematical formulas, complete with no more than 3 suggested keywords and JEL codes. We also encourage authors to make the introduction of their submitted articles understandable to the widest audience possible.

6. The first page of each submitted paper should contain every author's e-mail address, phone number and affiliation.

7. The editors or the publishing Society will not hold any responsibility for views expressed by authors in this Journal.

How to submit a paper

1. To submit a paper, at least one of the authors of the paper must be a member of the **Society for Promotion of Mechanism and Institution Design**. To join the Society, please go to the Society website: `http://www.mechanism-design.org/`.

   Currently, the annual membership rates are £60 (GBP) for regular members and £40 (GBP) for students. Concessions are available for members from low-income countries.

2. Papers should be submitted electronically to the **Journal of Mechanism and Institution Design** through the above website.

Aims & scope of the journal

The Journal of Mechanism and Institution Design aims to publish original articles that deal with the issues of designing, improving, analysing and testing economic, financial, political, or social mechanisms and institutions. It welcomes theoretical, empirical, experimental, historical and practical studies. It seeks creative, interesting, rigorous, and logical research and strives for clarity of thought and expression. We particularly encourage less experienced researchers such as recent PhD graduates to submit their work to the Journal and are sympathetic towards those papers that are novel and innovative but which, being outside the main paradigms, have been unsuccessful elsewhere. We hope that the published articles will be interesting and valuable to a broad audience from the areas of economics, finance, politics, law, computer science, management, history, mathematics, government, and related disciplines. The journal is an open-access, independent, peer-reviewed, non-profit, English-language journal with the purpose of disseminating and sharing the latest knowledge and understanding of the subject widely.

In order for any work that is published by the Journal of Mechanism and Institution Design to be freely accessible to the widest audience possible, when a paper is accepted by this Journal for publication, its author(s) will be asked to release the paper under the Creative Commons Attribution-Non-Commercial license. While the authors retain the copyright of their published work, this license permits anyone to copy and distribute the paper for non-commercial purposes provided that both the author(s) of the article and the Journal are properly acknowledged. For details of the copyrights, please see the "human-readable summary" of the license.