Incentive Compatibility and Feasibility Constraints in Housing Markets[*]

Antonio Nicolò[†]and Carmelo Rodríguez-Álvarez [‡]

October 7, 2011

Abstract

We analyze centralized housing markets under the existence of feasibility constraints on the number of agents and objects involved in the exchanges. We focus on an incomplete information setting where only the information about how each agent ranks her endowment is private. We show that under non-degenerate ex–ante probability distributions over preference profiles, no rule satisfies the joint requirements of individual rationality, (constrained) efficiency, and incentive compatibility.

JEL: C78; D02; D78.

Keywords: Housing Markets, Feasibility Constraints, Individual Rationality, k-Efficiency, Incentive compatibility.

---

[*] This article supersedes previous work contained on the working paper "Feasibility Constraints and Protective Behavior in Efficient Kidney Exchange".

[†] Dipartamento di Scienze Economiche "Marco Fanno". Universitá degli Studi di Padova. Via del Santo 33, 35123 PADOVA. Italy. `antonio.nicolo@unipd.it`. Fax: (+39) 049 827 4211. Phone: (+39) 049 827 4285.

[‡] Corresponding Author: Facultad CC. Económicas y Empresariales. Campus de Somosaguas. Universidad Complutense de Madrid. 28223 MADRID. Spain. `carmelor@ccee.ucm.es`. Fax: (+34) 91 394 2591. Phone: (+34) 91 394 2548

1 Introduction

A housing market problem consists of a set of agents who own a set indivisible objects (houses) that should be reallocated to the agents by a centralized clearinghouse. Each agent is entitled to keep the object she initially owns. Monetary transfers among the agents are not allowed. Real-life applications include paired kidney exchange, graduate housing, holiday house swaps, among others.[1] Since the seminal work by Shapley and Scarf (1974), housing markets have received considerable attention. When preferences are strict, there is a unique core outcome for each housing market in the sense that no group of agents can improve upon it by trading their initial endowments among them.[2] The core outcome can be obtained through the Gale's top trade cycle algorithm and the mechanism which selects it for each problem is the unique efficient mechanism that provides incentives for the agents to participate and to reveal their true preferences over objects.[3]

The classical framework of housing markets however, cannot be applied to some real-life problems of allocation of indivisible objects, because they have some specific features that have to be taken into consideration. In particular we are interested in problems where (i) the set of feasible allocations is restricted due to the presence of constraints on the number of exchanges that can be performed among the agents, and (ii) the domain of preferences is a subset of the universal domain. The presence of these two characteristics make these real-life problems interesting. Feasibility constraints render the top trade cycle algorithm inapplicable, because it may prescribe allocations that require arbitrarily large chains of exchange. The preference domain restriction raises the question whether there are other mechanisms which satisfy desirable properties and induce feasible allocations. Among the problems that share these features there are paired kidney exchange (PKE) programs and the holiday house swap services.

In PKE, there are logistic constraints in the number of donor-patient pairs involved in the exchange. In holiday house swaps, legal constraints may require exchanges to be pairwise. At the same time the central planner may have some information about agents' preferences which can be made public and induce a restricted preference domain. In PKE,

---

[1] See Sönmez and Ünver (2011) for a recent survey on the literature.

[2] See Shapley and Scarf (1974); Roth and Postlewaite (1977); Roth (1982).

[3] See Ma (1994).

patients' preferences over kidneys are based on compatibility and expected graft survival issues that are directly observable by doctors by means of medical tests. Preference profiles such that patients prefer incompatible organs to compatible ones can be excluded. In holiday house swap problems, there are characteristics of the houses that are desirables (or undesirable) for all the participants to the market and this is public information.

Nicolò and Rodríguez-Álvarez (2011b) consider this setting in a model of PKE and show that under feasibility constraints on the number of simultaneous kidney exchanges truthful revelation of patients' preferences is not compatible with a weak version of efficiency. Here, these negative results are extended proving that in a Bayesian setting, incentive compatible mechanisms do not satisfy that weak version of efficiency. In Nicolò and Rodríguez-Álvarez (2011a), positive results are restored when agents' preference have some degree of correlation.[4] In this paper we look to problems where there is a domain restriction induced by public information on some characteristics of the agents, but individual preferences can be totally uncorrelated. In particular, we assume that (i) how each agent ranks all pairs of objects different than her endowment is public information, (ii) there may exist a set of objects which are surely worse than the endowment for an agent. There are some objects however, whose ranking is common information and can be either better or worse than the initial endowment. In PKE, these assumptions capture the fact that doctors may have precise estimates of the expected graft survival of all available organs for each patient. Moreover, it is unambiguously accepted that remaining in dialysis is a better option than receiving an incompatible organ. These facts notwithstanding, the minimum expected graft survival such that each patient requires to undergo transplantation rather than waiting for a better match depends on unobservable characteristics like patient's risk aversion or time discounting. Consequently, this piece of relevant information remains private information of the patient.

The remainder of the paper is organized as follows. In Section 2, we present the model and basic notation. In Section 3, we introduce the notion of incentive compatibility and the impossibility theorem. In Section 4, we conclude with the proof of the theorem.

---

[4] In PKE, correlation of patients' preference can be observed because some characteristics of the donor, like age, affects in the same direction all patients. If two patients are compatible with the same pair of kidneys both prefer the organ of the youngest donor.

2 Housing Market Problems

Let $N = \{1, \ldots, n\}$ be a finite society consisting of a set of agents ($n \geq 3$). Each agent i owns an object ω_i, and let $\Omega = \{\omega_1, \ldots, \omega_n\}$ be the set of available objects.

A preference relation is a complete, reflexive, and transitive binary relation over Ω. Let $\mathcal{R}$ denote the set of all preference relations. Each agent $i \in N$ is equipped with a preference $\succsim_i \in \mathcal{R}$. For each $i \in N$ and each $\succsim_i \in \mathcal{R}$, $\succ_i$ is the associated strict preference, relation and by $\sim_i$ the associated indifference relation. A preference profile $\succsim$ is an n-tuples of preferences $\succsim = (\succsim_1, \ldots, \succsim_n)$, where for each $i \in N$, $\succsim_i \in \mathcal{R}$. For each $i \in N$ and each preference profile $\succsim$, $\succsim_{-i}$ is the complementary profile of preferences of all patients in $N \setminus \{i\}$.

For each agent i, her ranking of objects that are not ω_i is public information, but the position of ω_i in her preference is private information. For each $i \in N$, let $R_i \in \mathcal{R}$ be agent i's ***public preference***. A ***public preference profile*** $\mathbf{R}$ is an n-tuple of public preferences, $\mathbf{R} = (R_1, \ldots, R_n)$. Agent i's public preference does not completely defines i's preferences, but it determines her domain of admissible preferences. Specifically, for each $i \in N$, the preference $\succsim_i$ is consistent with the public preference profile $\mathbf{R}$ if for each pair $\omega, \omega' \in \Omega \setminus \{\omega_i\}$,

(i) $\omega \; R_i \; \omega' \Leftrightarrow \omega \succsim_i \omega'$,

(ii) $\omega_i \; R_i \; \omega$, implies $\omega_i \succ_i \omega$,

(iii) there is no $\omega \in \Omega \setminus \{\omega_i\}$ such that $\omega \sim_i \omega_0$.

By (i), i's preference over objects that are not ω_i coincide with R_i. By (ii), i's ranking of ω_i in $\succsim_i$ cannot be worse that its ranking according to R_i. By (iii), i is never indifferent between receiving any other agent's object or keeping ω_i. In words, two pieces of information may be public. First, how each agents ranks any pair of objects different than her own enodwment. Second, a set, if any, of objects which are surely worst than the endowment. Given $\mathbf{R}$, $\mathcal{D}_i^{\mathbf{R}}$ is the set of preferences which are consistent with $\mathbf{R}$ for agent i. Let $\mathcal{D}^{\mathbf{R}} \equiv \times_{i \in N} \mathcal{D}_i^{\mathbf{R}}$ and let $\mathcal{D}_{-i}^{\mathbf{R}} \equiv \times_{j \neq i} \mathcal{D}_j^{\mathbf{R}}$.

A (housing market) problem is a pair $(\mathbf{R}, \succsim)$ such that $\succsim \in \mathcal{D}^{\mathbf{R}}$.

An ***assignment*** is a bijection from objects to agents. We denote an arbitrary a as an n-tuple of pairs $a = [(1, \omega), \dots, (n, \omega')]$ such that for each $i, j \in N$, $i \neq j$ and each $\omega, \omega' \in \Omega$, if $(i, \omega), (j, \omega') \in a$, then $\omega \neq \omega'$. For each $i \in N$ and each assignment a, a_i is the object assigned to i according to a.

In every assignment, objects are allocated by forming exchange cycles. In each exchange cycle, every agent receives an object from some other agent in the cycle and simultaneously her object is assigned to another agent in the cycle. There are interesting applications of exchange of indivisible objects, like PKE, where logistic constraints restrict the number of agents involved in an exchange cycle. We incorporate such constraints in our analysis through a narrower definition of feasible assignments.

For each assignment a, let π_a be the finest partition of the set of agents such that for each $p \in \pi_a$ and each $i \in p$:

(i) either there are $j, j' \in p$, with $a_i = \omega_j$ and $a_{j'} = \omega_i$,[5]

(ii) or $a_i = \omega_0$.

Clearly, for each assignment a, the partition π_a is unique and well defined. We define the ***cardinality of*** a as the $max_{p \in \pi_a} \# p$.[6] The cardinality of an assignment refers to the size of the largest exchange cycle formed in the assignment.

For each $k \in \mathbb{N}$, $k \leq n$, we say that the assignment a is $k-$***feasible*** if a's cardinality is not larger than k. Let $\mathcal{A}^k$ be the set of all k–feasible assignments. An interesting case of feasibility restrictions appears when only immediate exchanges between two agents are admitted.

In this paper, we are interested in rules that select an assignment for each (exchange) problem. A ***rule*** is a mapping φ that selects an assignment a for each problem $(\mathbf{R}, \succsim)$. For each patient i and each problem $(\mathbf{R}, \succsim)$, we denote by $\varphi_i(\mathbf{R}, \succsim)$ the object assigned to i by φ at $(\mathbf{R}, \succsim)$. As $\mathbf{R}$ is given and public information, whenever there is no room for confusion, we drop $\mathbf{R}$ from the arguments and simply write $\varphi(\succsim)$. The assignment selected by a rule can be interpreted as an optimal recommendation that takes into account the

---

[5] Note that $j = j'$ and $i = j = j'$ are allowed.

[6] For each set S, $\#S$ refers to the number of elements of S.

preferences of the agents over objects and that tries to find a compromise between their (perhaps conflicting) interests.

We present a formal definition of the standard conditions for desirable rules. The reader should keep in mind that all the conditions refer to a given public preference profile $\mathbf{R}$.

Individual Rationality. For each $i \in N$ and each $\succsim \in \mathcal{D}^{\mathbf{R}}$, $\varphi_i(\succsim) \succsim_i \omega_0$.

$k-$***Efficiency.*** For each $\succsim \in \mathcal{D}^{\mathbf{R}}$, $\varphi(\succsim) \in \mathcal{A}^k$ and there is no $a \in \mathcal{A}^k$ such that for each $i \in N$ $a_i \succsim_i \varphi_i(\succsim)$ and for some $j \in N$, $a_j \succ_j \varphi_j(\succsim)$.

Individual rationality takes into account a patient's right to refuse any assignment that does not improve upon keeping her own object; *k-efficiency* is the natural version of efficiency that considers the feasibility restrictions on the cardinality of the assignments.

3 The Result

Agents preferences are private information and thus, they need to be elicited to propose an assignment. It is desirable that agents reveal their true preferences to ensure that the proposal is based on the correct information. This raises the problem of designing procedures that provide the right incentives for the agents to reveal truthfully this crucial information. Every $\mathbf{R}$ together with a rule $\varphi(\cdot)$ define a direct revelation mechanism. A direct revelation mechanism specifies a set of players (the patients), a set of strategies for each agent, the sets $\mathcal{D}_i^{\mathbf{R}}$, and an outcome function, $\varphi(\mathbf{R}, \cdot)$.[7] By the revelation principle, we focus without any loss of generality on agents' incentives under direct revelation mechanisms. Since we deal with probabilities, we need to consider agents' preferences for lotteries on assignments and we assume that agents are expected utility maximizers.

An ***expected utility function*** is a function $u : \Omega \to \mathbb{R}$. Given $\mathbf{R}$, for each $i \in N$ and each $\succsim_i \in \mathcal{D}_i^{\mathbf{R}}$, u_i is consistent with $\succsim_i$ if for each $\omega, \omega' \in \Omega$ such that $\omega \succsim_i \omega'$, $u_i(\omega) \geq u_i(\omega')$. An expected utility function reflects not only the simple order of her

---

[7] The direct revelation mechanism $(\mathbf{R}, \varphi)$ falls short of defining a game in strategic form because $\mathbf{R}$ does not introduce all the information about agents' preferences.

preferences regarding objects but also a measure of the intensity of her preferences. A ***lottery*** is a probability distribution over Ω. For an agent i with preferences $\succsim_i \in \mathcal{D}_i^{\mathbf{R}}$ and consistent utility function u_i, for each lottery λ, the utility of the lottery is given by its expected utility $\mathbb{E}[u_i \mid \lambda]$. For each pair of lotteries λ, λ', agent i (weakly) prefers λ to λ' if and only if $\mathbb{E}[u_i \mid \lambda] \geq \mathbb{E}[u_i \mid \lambda]$.

Let F denote a probability distribution over preference profiles. For each $i \in N$ and each $\succsim_i \in \mathcal{R}$, $\succsim_i$ is in the support of F if there is preference profile $\succsim^* \in \mathcal{R}$ such that $\mathrm{F}(\succsim^*) > 0$ and $\succsim_i^* = \succsim_i$. For each $i \in N$, each $\succsim_i \in \mathcal{R}$ in the support of F, we denote the marginal distribution over complementary profiles when agent i's preference is $\succsim_i$ by $\mathrm{F}(\cdot \mid \succsim_i)$. A probability distribution F is independent if for each $i \in N$ and each pair $\succsim_i, \succsim_i'$ in the support of F, $\mathrm{F}(\cdot \mid \succsim_i) = \mathrm{F}(\cdot \mid \succsim_i')$. For each independent probability distribution F, we denote i's marginal distribution over complementary profiles by F_i. Finally, abusing notation, for each $j \in N \setminus \{i\}$, and each $\succsim_j \in \mathcal{D}_j^{\mathbf{R}}$, and each independent probability distribution F, we define

$$\mathrm{f}_i(\succsim_j) \equiv \sum_{\{\succsim_{-i}^* \mid \succsim_j^* = \succsim_j\}} \mathrm{F}_i(\succsim_{-i}^*).$$

For a given $\mathbf{R}$, an ***information structure*** $\mathbf{F}$ is an independent probability distribution over preference profiles in $\mathcal{D}^{\mathbf{R}}$. For each $\succsim \in \mathcal{D}^{\mathbf{R}}$, $\mathrm{F}(\succsim)$ denotes the *ex-ante* probability that agents' preference profile is $\succsim$. Of course, for each $\succsim \in \mathcal{D}^{\mathbf{R}}$, $\mathrm{F}(\succsim) \geq 0$; for each $\succsim' \in \mathcal{R} \setminus \mathcal{D}^{\mathbf{R}}$, $\mathrm{F}(\succsim') = 0$; and $\sum_{\succsim \in \mathcal{D}^{\mathbf{R}}} \mathrm{F}(\succsim) = 1$.[8]

An information structure F is ***regular*** if for each preference profile $\succsim \in \mathcal{D}^{\mathbf{R}}$ such that for some $i \in N$ $\omega_i \succ_i \omega$ for each $\omega \in \Omega \setminus \{\omega_i\}$, $\mathrm{F}(\succsim) = 0$.

Regular information structures assign null probability to preference profiles where an agent considers her own object as the best preferred object and has no incentive to participate in the exchange.

---

[8] Note also that for each independent information structure F, for each $i \in N$, each $\succsim_i \in \mathcal{D}_i^{\mathbf{R}}$, and each pair $j, k \in N \setminus \{i\}$, $\mathrm{f}_j(\succsim_i) = \mathrm{f}_k(\succsim_i)$.

Incentive Compatibility. Given $\mathbf{R}$ and F, each agent $i \in N$, each pair $\succsim_i, \succsim_i' \in \mathcal{D}_i^{\mathbf{R}}$, and each u_i consistent with $\succsim_i$

$$\mathbb{E}\left[u_i(\varphi(\succsim_i, \cdot) \mid \mathrm{F}_i\right] \geq \mathbb{E}\left[u_i(\varphi(\succsim_i', \cdot) \mid \mathrm{F}_i\right].$$

Incentive compatibility requires that truth-telling for all patients forms a Bayesian Nash equilibrium of the Bayesian game defined by $(\mathbf{R}, \{u_i\}_{i\in N}, \varphi, F)$. Our main result is an impossibility theorem that extends the negative conclusions of Theorem 1 in Nicolò and Rodríguez-Álvarez (2011b) on the manipulability of rules in a PKE under complete information. In the incomplete information case, in equilibrium any revelation mechanism cannot satisfy the joint requirements of *k-efficiency*, and *individual rationality.*

Theorem 1. *Let* $2 \leq k < n$. *There exist public information profiles* $\mathbf{R}$ *such that for every regular information structure there is no rule that satisfies* individual rationality*,* k-efficiency*, and* incentive compatibility*.*

The impossibility result holds for arbitrary non-degenerate information structures.[9] Theorem 1 is a housing market counterpart to the main result in Roth (1989). This author shows that there is no incentive compatible mechanism that always selects a (core) stable outcome in two–sided one-to many matching markets.[10] In the context of PKE, Villa and Patrone (2009) show that a rule that maximizes the number of kidney exchanges does not satisfy *incentive compatibility* for some information structures. Theorem 1 considerably extends the result to any arbitrary rule and to all non-degenerate information structures.

4 Proofs

Proof of Theorem 1. Assume to the contrary there is a rule φ such that for each $\mathbf{R}$ there is a regular information structure F under which φ satisfies *individual rationality*, *k-*

---

[9] In fact, the proof of Theorem 1 applies to more general rules that depend also on the information structure.

[10] Roth (1989) deals with uncertainty about other agents' preferences (*private values.* Note that if a mechanism always selects a stable outcome, then it satisfies *efficiency* (in its range) and *individual rationality*). See Chakraborty et al. (2010) for related results in common value environment, where agents have uncertainty about their own preferences.

efficiency, and *incentive compatibility*. We consider the cases $k = 2$ and $2 < k < n$ separately.

Let $k = 2$. Consider three agents $\{1, 2, 3\}$ and public information profile $\mathbf{R}$ such that $\omega_2\ P_1\ \omega_3\ P_1\ \omega_1$, $\omega_3\ P_2\ \omega_1\ P_2\ \omega_2$, and $\omega_1\ P_3\ \omega_2\ P_3\ \omega_3$; and such that for each $i, j \in \{1, 2, 3\}$, $k \notin \{1, 2, 3\}$, $\omega_j\ P_i\ \omega_k$. Let $\succsim_1, \succsim_1' \in \mathcal{D}_1^{\mathbf{R}}$ be such that $\omega_2 \succ_1 \omega_3 \succ_1 \omega_1$ and $\omega_2 \succ_1' \omega_1 \succ_1' \omega_3$. Let $\succsim_2, \succsim_2' \in \mathcal{D}_2^{\mathbf{R}}$ be such that $\omega_3 \succ_2 \omega_1 \succ_2 \omega_2$ and $\omega_3 \succ_2' \omega_2 \succ_2' \omega_1$. Finally, let $\succsim_3 \in \mathcal{D}_3^{\mathbf{R}}$ be such that $\omega_1 \succ_3 \omega_2 \succ_3 \omega_3$ and $\omega_1 \succ_3' \omega_3 \succ_3' \omega_2$. Let F be such that for some $p, q, r \in (0, 1)$, $\mathrm{f}_i(\succsim_1) = p$, $\mathrm{f}_i\,(\succsim_1') = (1 - p)$ for $i \in \{2, 3\}$, $\mathrm{f}_j(\succsim_2) = q$, $\mathrm{f}_j(\succsim_2') = (1 - q)$ for $j \in \{1, 3\}$; and $\mathrm{f}_k(\succsim_3) = r$ and $\mathrm{f}_k(\succsim_3') = (1 - r)$ for $k \in \{1, 2\}$.

By *individual rationality* and *k-efficiency*,

$$\varphi(\succsim) \in \begin{cases} [(1, \omega_2), (2, \omega_1), (3, \omega_3)]\,, \\ [(1, \omega_3), (2, \omega_2), (3, \omega_1)]\,, \\ [(1, \omega_1), (2, \omega_3), (3, \omega_2)]\,. \end{cases}$$

Assume that $\varphi(\succsim) = [(1, \omega_2), (2, \omega_1), (3, \omega_3)]$. The arguments to prove the remaining cases are parallel since the problem is symmetric and F is arbitrary. By *individual rationality* and 2-*efficiency*,

$$\begin{aligned}
&\varphi(\succsim_1', \succsim_{-1}) \in \left\{ \begin{array}{l} [(1, \omega_2), (2, \omega_1), (3, \omega_3)]\,, \\ [(1, \omega_1), (2, \omega_3), (3, \omega_2)] \end{array} \right\}, \\
&\varphi(\succsim_2', \succsim_{-2}) \in \left\{ \begin{array}{l} [(1, \omega_3), (2, \omega_2), (3, \omega_1)]\,, \\ [(1, \omega_1), (2, \omega_3), (3, \omega_2)] \end{array} \right\}, \\
&\varphi(\succsim_3', \succsim_{-3}) \in \left\{ \begin{array}{l} [(1, \omega_3), (2, \omega_2), (3, \omega_1)]\,, \\ [(1, \omega_2), (2, \omega_1), (3, \omega_3)] \end{array} \right\}, \\
&\varphi(\succsim_1, \succsim_{-1}') = [(1, \omega_3), (2, \omega_2), (3, \omega_1)]\,, \\
&\varphi(\succsim_2, \succsim_{-2}') = [(1, \omega_2), (2, \omega_1), (3, \omega_3)]\,, \\
&\varphi(\succsim_3, \succsim_{-3}') = [(1, \omega_1), (2, \omega_3), (3, \omega_2)]\,, \\
&\varphi(\succsim') = [(1, \omega_1), (2, \omega_2), (3, \omega_3)]\,.
\end{aligned}$$

We explore eight mutually exclusive cases.

(i) $\varphi(\succsim_1', \succsim_{-1}) = [(1, \omega_2), (2, \omega_1), (3, \omega_3)]$, $\varphi(\succsim_2', \succsim_{-2}) = [(1, \omega_3), (2, \omega_2), (3, \omega_1)]$, and $\varphi(\succsim_3', \succsim_{-3}) = [(1, \omega_3), (2, \omega_2), (3, \omega_1)]$. Then, for each u_2 consistent with $\succsim_2$,

$$\begin{aligned}
&\mathbb{E}\left[u_2(\varphi(\succsim_2, \cdot) \mid \mathrm{F}_2\right] = \left[r + (1 - p)(1 - r)\right] u_2(\omega_1) + p(1 - r)u_2(\omega_2), \\
&\mathbb{E}\left[u_2(\varphi(\succsim_2'\ \cdot) \mid \mathrm{F}_2\right] = \left[(1 - r) + pr\right] u_2(\omega_2) + (1 - p)ru_2(\omega_3).
\end{aligned}$$

For each u_2 consistent with $\succsim_2$ such that $u_2(\omega_2) = 0$, $u_2(\omega_3) = 1$, and $u_2(\omega_1) < \frac{(1-p)r}{r+(1-p)(1-r)}$, $\mathbb{E}\left[u_2(\varphi(\succsim_2' \cdot) \mid \mathrm{F}_2\right] > \mathbb{E}\left[u_2(\varphi(\succsim_2, \cdot) \mid \mathrm{F}_2\right]$, which contradicts *incentive compatibility.*

(ii) $\varphi(\succsim_1', \succsim_{-1}) = [(1, \omega_2), (2, \omega_1), (3, \omega_3)]$, $\varphi(\succsim_2', \succsim_{-2}) = [(1, \omega_3), (2, \omega_2), (3, \omega_1)]$, and $\varphi(\succsim_3', \succsim_{-3}) = [(1, \omega_2), (2, \omega_1), (3, \omega_3)]$. Then, for each u_2 consistent with $\succsim_2$,

$$\begin{aligned} &\mathbb{E}\left[u_2(\varphi(\succsim_2, \cdot) \mid \mathrm{F}_2\right] = u_2(\omega_1), \\ &\mathbb{E}\left[u_2(\varphi(\succsim_2' \cdot) \mid \mathrm{F}_2\right] = \left[(1-r) + pr\right] u_2(\omega_2) + (1-p)r u_2(\omega_3). \end{aligned}$$

For each u_2 consistent with $\succsim_2$ such that $u_2(\omega_2) = 0$, $u_2(\omega_3) = 1$, and $u_2(\omega_1) < (1-p)r$, $\mathbb{E}\left[u_2(\varphi(\succsim_2' \cdot) \mid \mathrm{F}_2\right] > \mathbb{E}\left[u_2(\varphi(\succsim_2, \cdot) \mid \mathrm{F}_2\right]$, which contradicts *incentive compatibility.*

(iii) $\varphi(\succsim_1', \succsim_{-1}) = [(1, \omega_2), (2, \omega_1), (3, \omega_3)]$, $\varphi(\succsim_2', \succsim_{-2}) = [(1, \omega_1), (2, \omega_3), (3, \omega_2)]$, and $\varphi(\succsim_3', \succsim_{-3}) = [(1, \omega_3), (2, \omega_2), (3, \omega_1)]$. Then, for each u_2 consistent with $\succsim_2$,

$$\begin{aligned} &\mathbb{E}\left[u_2(\varphi(\succsim_2, \cdot) \mid \mathrm{F}_2\right] = \left[r + (1-p)(1-r)\right] u_2(\omega_1) + p(1-r)u_2(\omega_2), \\ &\mathbb{E}\left[u_2(\varphi(\succsim_2' \cdot) \mid \mathrm{F}_2\right] = (1-r)u_2(\omega_2) + r u_2(\omega_3). \end{aligned}$$

For each u_2 consistent with $\succsim_2$ such that $u_2(\omega_2) = 0$, $u_2(\omega_3) = 1$, and $u_2(\omega_1) < \frac{r}{r+(1-p)(1-r)}$, $\mathbb{E}\left[u_2(\varphi(\succsim_2' \cdot) \mid \mathrm{F}_2\right] > \mathbb{E}\left[u_2(\varphi(\succsim_2, \cdot) \mid \mathrm{F}_2\right]$, which contradicts *incentive compatibility.*

(iv) $\varphi(\succsim_1', \succsim_{-1}) = [(1, \omega_2), (2, \omega_1), (3, \omega_3)]$, $\varphi(\succsim_2', \succsim_{-2}) = [(1, \omega_3), (2, \omega_2), (3, \omega_1)]$, and $\varphi(\succsim_3', \succsim_{-3}) = [(1, \omega_2), (2, \omega_1), (3, \omega_3)]$. Then, for each u_2 consistent with $\succsim_2$,

$$\begin{aligned} &\mathbb{E}\left[u_2(\varphi(\succsim_2 \cdot) \mid \mathrm{F}_2\right] = u_2(\omega_1), \\ &\mathbb{E}\left[u_2(\varphi(\succsim_2' \cdot) \mid \mathrm{F}_2\right] = (1 - r + pr)u_2(\omega_2) + (1-p)r u_2(\omega_3). \end{aligned}$$

For each u_2 consistent with $\succsim_2$ such that $u_2(\omega_2) = 0$, $u_2(\omega_3) = 1$, and $u_2(\omega_1) < (1-p)r$, $\mathbb{E}\left[u_2(\varphi(\succsim_2' \cdot) \mid \mathrm{F}_2\right] > \mathbb{E}\left[u_2(\varphi(\succsim_2, \cdot) \mid \mathrm{F}_2\right]$, which contradicts *incentive compatibility.*

(v) $\varphi(\succsim_1', \succsim_{-1}) = [(1, \omega_1), (2, \omega_3), (3, \omega_2)]$, $\varphi(\succsim_2', \succsim_{-2}) = [(1, \omega_1), (2, \omega_3), (3, \omega_2)]$, and $\varphi(\succsim_3', \succsim_{-3}) = [(1, \omega_3), (2, \omega_2), (3, \omega_1)]$. Then, for each u_2 consistent with $\succsim_2$,

$$\begin{aligned} &\mathbb{E}\left[u_2(\varphi(\succsim_2, \cdot) \mid \mathrm{F}_2\right] = \left[pr + (1-p)(1-r)\right] u_2(\omega_1) + p(1-r)u_2(\omega_2) + (1-p)r u_2(\omega_3), \\ &\mathbb{E}\left[u_2(\varphi(\succsim_2' \cdot) \mid \mathrm{F}_2\right] = (1-r)u_2(\omega_2) + r u_2(\omega_3). \end{aligned}$$

For each u_2 consistent with $\succsim_2$ such that $u_2(\omega_2) = 0$, $u_2(\omega_3) = 1$, and $u_2(\omega_1) < \frac{pr}{pr+(1-p)(1-r)}$, $\mathbb{E}\left[u_2(\varphi(\succsim_2' \cdot) \mid \mathrm{F}_2\right] > \mathbb{E}\left[u_2(\varphi(\succsim_2, \cdot) \mid \mathrm{F}_2\right]$, which contradicts *incentive compatibility.*

(vi) $\varphi(\succsim_1', \succsim_{-1}) = [(1, \omega_1), (2, \omega_3), (3, \omega_2)]$, $\varphi(\succsim_2', \succsim_{-2}) = [(1, \omega_1), (2, \omega_3), (3, \omega_2)]$, and $\varphi(\succsim_3', \succsim_{-3}) = [(1, \omega_3), (2, \omega_2), (3, \omega_1)]$. Then, for each u_3 consistent with $\succsim_3$,

$$\begin{aligned}
&\mathbb{E}\left[u_3(\varphi(\succsim_3, \cdot) \mid \mathrm{F}_3\right] = (1 - pq)u_3(\omega_2) + pq\, u_3(\omega_3),\\
&\mathbb{E}\left[u_3(\varphi(\succsim_3' \cdot) \mid \mathrm{F}_3\right] = p\, u_3(\omega_1) + (1 - p)u_3(\omega_3).
\end{aligned}$$

For each u_3 consistent with $\succsim_3$ such that $u_3(\omega_1) = 1$, $u_3(\omega_3) = 0$, and $u_3(\omega_2) < \frac{p}{1-pq}$, $\mathbb{E}\left[u_3(\varphi(\succsim_3' \cdot) \mid \mathrm{F}_3\right] > \mathbb{E}\left[u_3(\varphi(\succsim_3, \cdot) \mid \mathrm{F}_3\right]$, which contradicts *incentive compatibility.*

(vii) $\varphi(\succsim_1', \succsim_{-1}) = [(1, \omega_1), (2, \omega_3), (3, \omega_2)]$, $\varphi(\succsim_2', \succsim_{-2}) = [(1, \omega_1), (2, \omega_3), (3, \omega_2)]$, and $\varphi(\succsim_3', \succsim_{-3}) = [(1, \omega_2), (2, \omega_1), (3, \omega_3)]$. Then, for each u_3 consistent with $\succsim_3$,

$$\begin{aligned}
&\mathbb{E}\left[u_3(\varphi(\succsim_3, \cdot) \mid \mathrm{F}_3\right] = (1 - pq)u_3(\omega_2) + pq\, u_3(\omega_3),\\
&\mathbb{E}\left[u_3(\varphi(\succsim_3' \cdot) \mid \mathrm{F}_3\right] = p(1 - q)u_3(\omega_1) + (1 - p + pq)u_3(\omega_3).
\end{aligned}$$

For each u_3 consistent with $\succsim_3$ such that $u_3(\omega_1) = 1$, $u_3(\omega_3) = 0$, and $u_3(\omega_2) < \frac{p(1-q)}{1-pq}$, $\mathbb{E}\left[u_3(\varphi(\succsim_3' \cdot) \mid \mathrm{F}_3\right] > \mathbb{E}\left[u_3(\varphi(\succsim_3, \cdot) \mid \mathrm{F}_3\right]$, which contradicts *incentive compatibility.*

(viii) $\varphi(\succsim_1', \succsim_{-1}) = [(1, \omega_1), (2, \omega_3), (3, \omega_2)]$, $\varphi(\succsim_2', \succsim_{-2}) = [(1, \omega_3), (2, \omega_2), (3, \omega_1)]$, and $\varphi(\succsim_3', \succsim_{-3}) = [(1, \omega_2), (2, \omega_1), (3, \omega_3)]$. Then, for each u_1' consistent with $\succsim_1'$,

$$\begin{aligned}
&\mathbb{E}\left[u_1'(\varphi(\succsim_1, \cdot) \mid \mathrm{F}_3\right] = qu_1'(\omega_2) + (1 - q)u_1'(\omega_3),\\
&\mathbb{E}\left[u_1'(\varphi(\succsim_1' \cdot) \mid \mathrm{F}_1\right] = \left[qr + (1 - q)\right] u_1'(\omega_1) + q(1 - r)u_1'(\omega_2).
\end{aligned}$$

For each u_1' consistent with $\succsim_1'$ such that $u_1'(\omega_1) = 0$, $u_1'(\omega_2) = 1$, and $u_1'(\omega_3 > \frac{-qr}{1-q}$, $\mathbb{E}\left[u_1'(\varphi(\succsim_1 \cdot) \mid \mathrm{F}_1\right] > \mathbb{E}\left[u_1'(\varphi(\succsim_1', \cdot) \mid \mathrm{F}_1\right]$, which contradicts *incentive compatibility.*

Next, we analyze the case $2 < k < n$. Remember that there are at least $k + 1$ agents. Let $\mathbf{R}$ be such that for each $i = 1 \ldots, k + 1$:

$$\begin{array}{ll}
\omega_{i+1}\ P_i\ \omega_i\ P_i\ \omega_j & \forall i \notin \{k-1, k+1\},\ \forall j \notin \{i, i+1\},\\
\omega_k\ P_{k-1}\ \omega_{k+1}\ P_{k-1}\ \omega_{k-1}\ P_{k-1}\ \omega_j' & \forall j' \notin \{k-1, k, k+1\},\\
\omega_1\ P_{k+1}\ \omega_2\ P_{k+1}\ \omega_{k+1}\ P_{k+1}\ \omega_j'' & \forall j'' \notin \{1, 2, k+1\}.
\end{array}$$

By *individual rationality*, without any loss of generality, we can focus on the assignment restricted to the patients $1,\dots,k+1$.

For each $i \in N$, let $\succsim_i = P_i$, $\succsim'_{k-1} \in \mathcal{R}^{\mathbf{P}}_{k-1}$ be such that $\omega_k \succ'_{k-1} \omega_{k-1} \succ'_{k-1} \omega_{k+1}$ and $\succsim'_{k+1} \in \mathcal{R}^{\mathbf{P}}_{k+1}$ be such that $\omega_1 \succ'_{k+1} \omega_{k+1} \succ'_{k+1} \omega_2$. Let F be such that for each $i \in N \backslash \{k-1, k+1\}$ and $j \in \{k-1, k+1\}$, $\mathrm{f}_j(\succsim_i) = 1$, $\mathrm{f}_{k+1}(\succsim_{k-1}) = p$, $\mathrm{f}_{k+1}(\succsim'_{k-1}) = (1-p)$, $\mathrm{f}_{k-1}(\succsim_{k+1}) = q$, and $\mathrm{f}_{k-1}(\succsim'_{k+1}) = (1-q)$. By *individual rationality* and *k-efficiency*,

$$\varphi(\succsim) \in \left\{ \left[\begin{array}{ll} (1,\omega_1) & \\ (i,\omega_{i+1}) & \forall i = 2,\dots,k \\ (k+1,\omega_2) & \end{array}\right], \left[\begin{array}{ll} (i,\omega_{i+1}) & \forall i < k-1 \\ (k-1,\omega_{k+1}) & \\ (k,\omega_k) & \\ (k+1,\omega_1) & \end{array}\right] \right\},$$

$$\varphi(\succsim'_{k-1}, \succsim_{-(k-1)}) = \left[\begin{array}{ll} (1,\omega_1) & \\ (i,\omega_{i+1}) & \forall i = 2,\dots,k \\ (k+1,\omega_2) & \end{array}\right],$$

$$\varphi(\succsim'_{k+1}, \succsim_{-(k+1)}) = \left[\begin{array}{ll} (i,\omega_{i+1}) & \forall i < k-1 \\ (k-1,\omega_{k+1}) & \\ (k,\omega_k) & \\ (k+1,\omega_1) & \end{array}\right],$$

$$\varphi(\succsim'_{k+1}, \succsim'_{k+1}, \succsim_{-(\{k-1,k+1\})}) = [(i,\omega_i)\ \forall i \in N].$$

Consider agent $k-1$. If $\varphi_{k-1}(\succsim) = \omega_{k+1}$, then for each u_{k-1} consistent with $\succsim'_{k-1}$,

$$\begin{aligned} &\mathbb{E}\left[u_{k-1}(\varphi(\succsim_{k-1},\cdot) \mid \mathrm{F}_{k-1}\right] = u_{k-1}(\omega_{k+1}), \\ &\mathbb{E}\left[u_{k-1}(\varphi(\succsim'_{k-1}\ \cdot) \mid \mathrm{F}_{k-1}\right] = q u_{k-1}(\omega_k) + (1-q) u_{k-1}(\omega_{k-1}), \end{aligned}$$

For each u_{k-1} consistent with $\succsim_{k-1}$, such that $u_{k-1}(\omega_{k-1}) = 0$, and $u_{k-1}(\omega_k) > \frac{1}{q} u_{k-1}(\omega_{k+1})$, $\mathbb{E}\left[u_{k-1}(\varphi(\succsim'_{k-1}\ \cdot) \mid \mathrm{F}_{k-1}\right] > \mathbb{E}\left[u_{k-1}(\varphi(\succsim_{k-1},\cdot) \mid \mathrm{F}_{k-1}\right]$, which contradicts *incentive compatibility*. Then, $\varphi_{k-1}(\succsim) = \omega_k$ and

$$\varphi(\succsim) = \left[\begin{array}{ll} (1,\omega_1) & \\ (i,\omega_{i+1}) & \forall i = 2,\dots,k \\ (k+1,\omega_2) & \end{array}\right].$$

Note that for each u_{k+1} consistent with $\succsim_{k+1}$,

$$\begin{aligned} &\mathbb{E}\left[u_{k+1}(\varphi(\succsim_{k+1},\cdot) \mid \mathrm{F}_{k+1}\right] = u_{k+1}(\omega_2), \\ &\mathbb{E}\left[u_{k+1}(\varphi(\succsim'_{k+1}\ \cdot) \mid \mathrm{F}_{k+1}\right] = p u_{k+1}(\omega_1) + (1-p) u_{(}\omega_{k+1}), \end{aligned}$$

For each u_{k+1} consistent with $\succsim_{k+1}$, such that $u_{k+1}(\omega_{k+1}) = 0$, and $u_{k+1}(\omega_1) > \frac{1}{p}u_{k+1}(\omega_2)$, $\mathbb{E}\left[u_{k+1}(\varphi(\succsim'_{k+1}\ \cdot) \mid \mathrm{F}_{k+1}\right] > \mathbb{E}\left[u_{k+1}(\varphi(\succsim_{k+1}, \cdot) \mid \mathrm{F}_{k+1}\right]$, which contradicts *incentive compatibility.* ☐

References

A. Chakraborty, A. Cittana, and M. Ostrovsky. Two–sided matching with interdependent values. *Journal of Economic Theory*, 145(1):85–105, 2010.

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

A. Nicolò and C. Rodríguez-Álvarez. Age-based preferences: Incorporating compatible pairs into paired kidney exchange. 2011a. Unpublished manuscript, Universitá degli Studi de Padova and Universidad Complutense de Madrid.

A. Nicolò and C. Rodríguez-Álvarez. Transplant quality and patients' preferences in paired kidney exchange. *Games and Economic Behavior*, In Press, 2011b. doi: DOI: 10.1016/j.geb.2011.05.009.

A. E. Roth. Incentive compatibility in a market with indivisible goods. *Economics Letters*, 9:127–132, 1982.

A. E. Roth. Two–sided matching with incomplete information about others' preferences. *Games and Economic Behavior*, 1:191–209, 1989.

A. E. Roth and A. Postlewaite. Weak versus strong domination in a market of indivisible goods. *Journal of Mathematical Economics*, 4:131–137, 1977.

L. Shapley and H. Scarf. On cores and indivisibility. *Journal of Mathematical Economics*, 1:23–37, 1974.

T. Sönmez and M. U. Ünver. Matching, allocation, and exchange of discrete resources. In *Handbook of Social Economics,* Eds. Jesse Benhabib and Alberto Bisin and Matthew O. Jackson. North Holland, The Netherlands, pages 751–782, 2011.

S. Villa and F. Patrone. Incentive compatibility in kidney exchange problems. *Health Care Management Science*, 12:351–362, 2009.