Check for updates

ORIGINAL ARTICLE

International Journal of Economic Theory

JJET WILEY

Preference revelation games and strict cores of multiple-type housing market problems

Di Feng | Bettina Klaus

Faculty of Business and Economics, University of Lausanne, Lausanne, Switzerland

Correspondence

Bettina Klaus, Faculty of Business and Economics, University of Lausanne, 1015 Lausanne, Switzerland. Email: Bettina.Klaus@unil.ch

Abstract

We consider multiple-type housing market problems as introduced by Moulin (1995) and study the relationship between strict strong Nash equilibria and the strict core (two solution concepts that are defined in terms of the absence of weak blocking coalitions). We prove that for lexicographically separable preferences, the set of all strict strong Nash equilibrium outcomes of each preference revelation game that is induced by a strictly core stable mechanism is a subset of the strict core, but not vice versa, that is, there are strict core allocations that cannot be implemented in strict strong Nash equilibrium. This result is extended to a more general set of preference domains that satisfy strict core nonemptiness and a minimal preference domain richness assumption.

KEYWORDS

multiple-type housing market problems, strict core, strict strong Nash equilibria

JEL CLASSIFICATION C71, C72, C78

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. International Journal of Economic Theory published by John Wiley & Sons Australia, Ltd on behalf of International Association for Economic Theory.

1 | INTRODUCTION

WILEY-

In a classical Shapley–Scarf housing market (Shapley and Scarf, 1974), each agent is endowed with an indivisible object, such as a house, wishes to consume exactly one house, and ranks all houses in the market. The problem then is to (re)allocate houses among the agents without using monetary transfers and by taking into account agents' preferences and endowments.

A common solution concept for Shapley–Scarf housing markets is the strict core solution, which assigns the set of allocations where no group of agents has an incentive (via weak blocking) to deviate by exchanging their endowments within the group. When agents' preferences are strict, the strict core solution exhibits a remarkable number of positive features: it is non-empty (Shapley and Scarf, 1974), always a singleton, and coincides with the unique competitive allocation (Roth and Postlewaite, 1977). In addition, it can be easily calculated by the so-called top trading cycles (TTC) algorithm (due to David Gale). Furthermore, the TTC mechanism that assigns the unique strict core allocation for any housing market is strategy-proof (Roth, 1982), and it is the unique mechanism satisfying individual rationality, Pareto efficiency, and strategy-proofness (Ma, 1994).

Multiple-type housing markets are an extension of Shapley–Scarf housing markets, which were first introduced by Moulin (1995).¹ In multiple-type housing markets, there are multiple types of indivisible objects, each agent is endowed with one object of each type and wishes to consume exactly one object of each type. Multiple-type housing markets are often described with houses and cars as metaphors for indivisible object types. While these and related housing market models appear to be rather stylized, they give valuable insights into many real-world applications such as dynamic resource allocation problems (Monte and Tumennasan, 2015), the assignment of student presentations (Mackin and Xia, 2016), and the assignment of medical resources (Huh et al., 2013). A more familiar example for most readers would be the situation of students' enrollment at many universities where courses are taught in parallel sessions (Klaus, 2008).

Konishi et al. (2001) were the first to analyze multiple-type housing markets. They demonstrated that when increasing the dimension of the classical Shapley–Scarf housing market model by adding other types of indivisible objects, most of the positive results obtained for the one-dimensional single-type case disappear: even for additively separable preferences, the strict core may be empty and no individually rational, Pareto efficient, and strategy-proof mechanism exists. One of the reasons for this is that, in contrast to single-type housing market problems, multiple-type housing market problems cannot be transformed into well-behaved coalition formation games (Banerjee et al., 2001; Bogomolnaia and Jackson, 2002; Quint and Wako, 2004); for example, an agent may exchange his house within a trading coalition S_1 but exchange his car with a different trading coalition S_2 .

There has been very little work on multiple-type housing market problems since Konishi et al.'s (2001) negative results. The following papers considered different solutions for different sub-domains of preferences.

For separable preferences, Konishi et al. (2001) and Wako (2005) suggested an alternative solution to the strict core solution by first using separability to decompose a multiple-type housing market into "coordinatewise submarkets" and then determining the strict core in each submarket. Wako (2005) called the resulting outcome the commoditywise competitive allocation and showed that it is implementable in strong Nash equilibria. Klaus (2008) called the

2

¹There are many other extensions, such as the multi-demand models of Pápai (2001, 2007), Ehlers and Klaus (2003), and Manjunath and Westkamp (2021).

IJET -WILEY 3

mechanism that always assigns this unique allocation the coordinatewise core rule, and showed that it satisfies individual rationality, constrained efficiency,² and strategy-proofness.

For a very general domain of lexicographic preferences, Sikdar et al. (2017, 2019) extended the TTC algorithm and defined a new mechanism: the multiple-type top trading cycles (MTTC) mechanism, and they showed that the MTTC mechanism determines a strict core allocation; hence, the strict core for general lexicographic preferences is non-empty. Strict core stability implies individual rationality and Pareto efficiency of the MTTC mechanism. However, they demonstrated that the MTTC mechanism is not strategy-proof and that the strict core may be multi-valued.

1.1 | Our contributions

Takamiya (2009) considered the more generalized model of indivisible goods allocation introduced by Sönmez (1999) that contains Shapley–Scarf housing market problems as special case. In particular, Takamiya's results imply that for Shapley–Scarf housing market problems and for individually rational and Pareto efficient mechanisms, the set of strict strong Nash equilibrium outcomes of the preference revelation game equals the strict core (we state this result as Corollary 1).

Similarly, we examine the relationship between strict strong Nash equilibrium outcomes of the preference revelation games and strict core allocations for multiple-type housing markets. Takamiya's (2009) results do not translate into our higher-dimensional model. First, multiple-type housing market problems may have an empty strict core, even if preferences are separable. Then, a promising subdomain that guarantees the non-emptiness of the strict core is the domain of lexicographically separable preferences. However, lexicographically separable preferences do not satisfy the domain richness condition Takamiya (2009) needs for his main result.

We prove that, for lexicographically separable preferences, the set of all strict strong Nash equilibrium outcomes of each preference revelation game that is induced by a strictly core stable mechanism is a subset of the strict core, but not vice versa, that is, there are strict core allocations that cannot be implemented in strict strong Nash equilibrium (Theorem 1). This result is extended to a more general set of preference domains that satisfy strict core non-emptiness and a minimal preference domain richness assumption (Theorem 2). Throughout the paper, we motivate our approach and discuss some comparative statics aspects of our results via various examples.

2 | THE MODEL

2.1 | Multiple-type housing market problems

Let $N = \{1, ..., n\}$ be a finite set of agents. A non-empty subset of agents $S \subseteq N$ is a coalition. We assume that there exist $m \ge 1$ (distinct) types of indivisible objects and n (distinct) indivisible objects of each type. We denote the set of types by $T = \{1, ..., m\}$. Note that for m = 1, our model equals the classical Shapley–Scarf housing market model (Shapley and Scarf, 1974).

²There exists no other strategy-proof mechanism that Pareto dominates the coordinatewise core rule.

Throughout this paper, we focus on the multiple-type extension of the Shapley–Scarf housing market model as introduced by Moulin (1995), where $|N| = n \ge 3$ and $|T| = m \ge 2$.³

Each agent $i \in N$ is endowed with exactly one object of each type $t \in T$, denoted by o_i^t . Hence, each agent *i*'s endowment is a list $o_i = (o_i^1, ..., o_i^m)$. The set of type-t objects is $O^t = \{o_1^t, ..., o_n^t\}$, and the set of all objects is $O = \{o_1^1, o_1^2, ..., o_n^m, ..., o_n^n\}$. In particular, $|O| = n \times m$.

For each agent *i*, an allotment x_i assigns one object of each type to agent *i*, that is, x_i is a list $x_i = (x_i^1, ..., x_i^m) \in \prod_{t \in T} O^t$, where $x_i^t \in O^t$ is agent i's type-t allotment. Alternatively, we sometimes denote an allotment as a subset of objects, $x_i = \{x_i^1, ..., x_i^m\} \subsetneq O$, and refer to a subset of an allotment as a partial allotment. We assume that each agent *i* has complete, antisymmetric, and transitive preferences R_i over all possible allotments, that is, R_i is a linear order over $\prod_{t \in T} O^t$. For two allotments x_i and y_i, x_i is weakly better than y_i if $x_i R_i y_i$, and x_i is strictly better than y_i if $[x_i R_i y_i$ and not $y_i R_i x_i]$, denoted $x_i P_i y_i$. Finally, since preferences over allotments are strict, x_i is indifferent to y_i only if $x_i = y_i$. We denote preferences as ordered lists, for example, $R_i : x_i, y_i, z_i$ instead of $x_i P_i y_i P_i z_i$. The set of all preferences is denoted by \mathcal{R} , which we will also refer to as the strict preference domain.

A preference profile specifies preferences for all agents and is denoted by a list $R = (R_1, ..., R_n) \in \mathcal{R}^N$. We use the following standard notation $R_{-i} = (R_1, ..., R_{i-1}, R_{i+1}, ..., R_n)$ to denote the list of all agents' preferences, except for agent *i*'s preferences. Furthermore, for each coalition *S* we define $R_S = (R_i)_{i \in S}$ and $R_{-S} = (R_i)_{i \in N \setminus S}$ to be the lists of preferences of coalitions *S* and *N**S*, respectively.

In addition to the domain of strict preferences, we are considering several preference subdomains based on agents' "marginal preferences": assume that for each agent $i \in N$ and for each type $t \in T$, *i* has complete, antisymmetric, and transitive preferences R_i^t over the set of type-*t* objects O^t . We refer to R_i^t as agent *i*'s type-*t* marginal preferences, and denote by \mathcal{R}^t the set of all type-*t* marginal preferences. Then we can define the following two preference domains.

Definition 1 (Separability). Agent *i*'s preferences $R_i \in \mathcal{R}$ are *separable* if for each $t \in T$ there exist type-*t* marginal preferences $R_i^t \in \mathcal{R}^t$ such that for any two allotments x_i and y_i , if for all $t \in T$, $x_i^t R_i^t y_i^t$, then $x_i R_i y_i$. \mathcal{R}_s denotes the *domain of separable preferences*.

Before turning to our next preference domain, we introduce some notation. We use a bijective function $\pi_i : T \to T$ to order types according to agent *i*'s "(subjective) importance," with $\pi_i(1)$ being the most important and $\pi_i(m)$ being the least important object type. We denote by π_i an ordered list of types; for example, by $\pi_i = (2, 3, 1)$ we mean that $\pi_i(1) = 2, \pi_i(2) = 3$, and $\pi_i(3) = 1$. So for each agent $i \in N$ and each allotment $x_i = (x_i^1, ..., x_i^m)$, by $x_i^{\pi_i} = (x_i^{\pi_i(1)}, ..., x_i^{\pi_i(m)})$ we denote the allotment after rearranging it with respect to the *object-type importance order* π_i .

Definition 2 (Lexicographic separability). Agent *i*'s preferences $R_i \in \mathcal{R}$ are *lexicographically* separable if they are separable with type-*t* marginal preferences $(R_i^t)_{t \in T}$ and there exists an object-type importance order $\pi_i : T \to T$ such that for any two allotments x_i and y_i , if $x_i^{\pi_i(1)}P_i^{\pi_i(1)}y_i^{\pi_i(1)}$ or if there exists a positive integer $k \leq m - 1$ such that $x_i^{\pi_i(1)} =$

³One-agent and two-agent multiple-type housing market problems are rather trivial cases due to the fact that no trade or pairwise trade are the only possibilities. The real complexity of the model, however, arises from the possibility that agents are simultaneously trading in different coalitions.

 $y_i^{\pi_i(1)}$, ..., $x_i^{\pi_i(k)} = y_i^{\pi_i(k)}$, and $x_i^{\pi_i(k+1)}P_i^{\pi_i(k+1)}y_i^{\pi_i(k+1)}$, then $x_iP_iy_i$. \mathcal{R}_l denotes the domain of *lexicographically separable preferences*.

Remark 1 (Representation of lexicographically separable preferences). Note that any lexicographically separable preference relation $R_i \in \mathcal{R}_l$ is uniquely determined by agent *i*'s marginal preferences $(R_i^l)_{l \in T}$ and an object-type importance order π_i . For example, consider a situation with $T = \{H(ouse), C(ar)\}$ and $N = \{1, 2, 3\}$ with each agent *i*'s endowment equal to $o_i = (H_i, C_i)$. Assume that agent *i* has lexicographically separable preferences $R_i : (H_1, C_1), (H_1, C_2), (H_1, C_3), (H_2, C_1), (H_2, C_2), (H_2, C_3), (H_3, C_1), (H_3, C_2), (H_3, C_3)$. Then, agent *i*'s type importance order is $\pi_i : H, C$, and his marginal preferences are $R_i^H : H_1, H_2, H_3$, and $R_i^C : C_1, C_2, C_3$. Hence, agent *i*'s preferences R_i can alternatively be written as $R_i = (R_i^H, R_i^C, \pi_i)$.

For an even more compact description of agent *i*'s lexicographically separable preferences, we can also rely on the strict ordering of objects that is induced by the object-type importance order together with his marginal preferences: $\mathbf{R}_i : H_1, H_2, H_3, C_1, C_2, C_3$.

An *allocation* x partitions the set of all objects O into agents' allotments, that is, $x = \{x_1, ..., x_n\}$ is such that for each $t \in T$, $\bigcup_{i \in N} x_i^t = O^t$ and, for each pair $i \neq j, x_i^t \neq x_j^t$. For simplicity, sometimes we will restate an allocation as a list $x = (x_1, ..., x_n)$. The *set of all allocations* is denoted by X, and the *endowment allocation* is denoted by $e = (o_1, ..., o_n)$. Given x, we define $x_{-i} = (x_1, ..., x_{i-1}, x_{i+1}, ..., x_n)$ to be the list of allotments of all agents except for agent *i*'s allotment and $x_S = (x_i)_{i \in S}$ to be the list of allotments of coalition S.

We assume that when facing an allocation x, there are no consumption externalities and each agent $i \in N$ only cares about his own allotment x_i . Hence, each agent i's preferences over allocations X are essentially equivalent to his preferences over allotments $\Pi_{t \in T} O^t$. With some abuse of notation, we use notation R_i to denote an agent i's preferences over allotments as well as his preferences over allocations, that is, for each agent $i \in N$ and for any two allocations $x, y \in X, x R_i y$ if and only if $x_i R_i y_i$.⁴

A (*multiple-type housing market*) problem is a triple (N, e, R); as the set of agents N and the endowment allocation e remain fixed throughout, we will simply denote problem (N, e, R) by R. Thus, the strict preference profile domain \mathcal{R}^N also denotes the set of all problems.

2.2 | Solutions/mechanisms and their properties

Note that all following definitions for the domain of strict preferences \mathcal{R} can alternatively be formulated for any subdomain $\hat{\mathcal{R}} \subseteq \mathcal{R}$.

A solution is a set-valued function $F : \mathcal{R}^N \to 2^X$ that assigns to each problem $R \in \mathcal{R}^N$ a (possibly empty) set of allocations $F(R) \subseteq X$. A mechanism is a function $f : \mathcal{R}^N \to X$ that assigns to each problem $R \in \mathcal{R}^N$ an allocation $f(R) \in X$, and for each $i \in N, f_i(R)$ is agent i's allotment under mechanism f at R.

We next introduce and discuss some well-known properties for allocations, solutions, and mechanisms. First we consider a voluntary participation condition for an allocation x to be

⁴Note that when extending strict preferences over allotments to preferences over allocations without consumption externalities, strictness is lost because any two allocations where an agent gets the same allotment are indifferent to that agent.

implementable without causing agents any harm: no agent will be worse off than at his endowment.

Definition 3 (Individual rationality). An allocation $x \in X$ is *individually rational* if for each agent $i \in N$, $x_i R_i o_i$. A solution/mechanism is *individually rational* if for each problem $R \in \mathcal{R}^N$, it assigns only individually rational allocations.

Next, we consider a well-known efficiency criterion.

Definition 4 (Pareto efficiency). An allocation $y \in X$ Pareto dominates allocation $x \in X$ if for each agent $i \in N$, $y_i R_i x_i$, and for at least one agent $j \in N$, $y_i P_i x_i$. An allocation $x \in X$ is *Pareto efficient* if there is no allocation $y \in X$ that Pareto dominates it. A solution/mechanism is *Pareto efficient* if for each problem $R \in \mathcal{R}^N$, it assigns only Pareto efficient allocations.

Next, we define an incentive property for mechanisms that models that no agent can benefit from misrepresenting his preferences.

Definition 5 (Strategy-proofness). A mechanism f is strategy-proof if for each problem $R \in \mathcal{R}^N$, each agent $i \in N$, and each preference relation $R'_i \in \mathcal{R}$, we have $f_i(R_i, R_{-i})R_i f_i(R_i', R_{-i}).$

Next, in order to introduce the standard cooperative solutions of the weak and the strict core, we introduce two blocking notions: for problem $R \in \mathcal{R}^N$, an allocation $x \in X$ is strictly *blocked by coalition* $S \subseteq N$ if there exists an allocation $y \in X$ such that

- (1) at allocation y agents in S reallocate their endowments, that is, for each $i \in S$ and each $t \in T, y_i^t \in \prod_{j \in S} o_j^t$; and
- (2) all agents in S are strictly better off, that is, for each $i \in S$, $y_i P_i x_i$.

An allocation $x \in X$ is weakly blocked by coalition $S \subseteq N$ if there exists an allocation $y \in X$ such that (1) and

(2') all agents in S are weakly better off, with at least one of them being strictly better off, that is, for each $i \in S$, $y_i R_i x_i$, and for some $j \in S$, $y_i P_j x_j$.

Given the blocking notions above, we can restate individual rationality and Pareto efficiency as follows. An allocation is individually rational if it is not weakly or strictly blocked by any singleton coalition $\{i\}$ and an allocation is Pareto efficient if it is not weakly blocked by the set of all agents N.

We now introduce the first type of (possibly empty- or multi-valued) solution to multipletype housing market problems that we will consider: core solutions.

Definition 6 (Strict/weak core-stability). An allocation is a *strict/weak core allocation* if it is not weakly/strictly blocked by any coalition; the set of all strict/weak core allocations is the *strict/weak core*. Given a problem $R \in \mathcal{R}^N$, let SC(R)/WC(R) denote its strict/weak core. A mechanism f is strictly/weakly core stable if for any problem, it assigns only strict/ weak core allocations.

Note that for all problems $R \in \mathcal{R}^N$, $SC(R) \subseteq WC(R)$, and that all strict core allocations satisfy individual rationality and Pareto efficiency. So, if a mechanism is strictly core stable, then it is individually rational and Pareto efficient as well. Furthermore, for some problems $R \in \mathcal{R}^N$, WC(R) may be empty.⁵

We next focus on the domain of lexicographically separable preferences (\mathcal{R}_l) and extend Gale's famous TTC algorithm to multiple-type housing market problems. More specially, we adapt the MTTC introduced by Sikdar et al. (2017, 2019) to \mathcal{R}_l .⁶

Definition 7 (The multiple-type top trading cycles [MTTC] algorithm/mechanism).

Input. A multiple-type housing market problem $R \in \mathcal{R}_l^N$.

Step 1. Building step. Let N(1) = N and U(1) = O. We construct a directed graph G(1) with the set of nodes $N(1) \cup U(1)$. For each $o \in U(1)$, we add an edge from the object to its owner and for each $i \in N(1)$, we add an edge from the agent to his most preferred object in O (according to the linear representation of R_i we explained in Remark 1). For each edge $(i, o) \in N \times O$ we say that agent *i* points to object *o*.

Implementation step. A *trading cycle* is a directed cycle in graph G(1). Given the finite number of nodes, at least one trading cycle exists. We assign to each agent *i* in a trading cycle the object that he pointed to, and denote the object assigned to him in this step by $a_i(1)$; we denote the corresponding set of objects assigned through trading cycles by A(1). If agent $i \in N$ was part of a trading cycle, then his partial allotment $x_i(1) = \{a_i(1)\}$; otherwise, $x_i(1) = \emptyset$.

Removal step. We remove all objects that were assigned through trading cycles from set O and set $U(2) := O \setminus A(1)$, which are the objects that have not yet been allocated. For each agent $i \in N$, we derive the set of *feasible continuation objects* $U_i(2)$ by removing all objects in U(2) that are of a type that is already present in agent *i*'s partial allotment $x_i(1)$. Since $m \ge 2$, no agents are removed in this step and we let N(2) := N. Go to step 2.

In general, at step $q (\geq 2)$ we have the following.

Step q. If U(q) (or equivalently N(q)) is empty, then stop; otherwise do the following.

Building step. We construct a directed graph G(q) with the set of nodes $N(q) \cup U(q)$. For each $o \in U(q)$, we add an edge from the object to its owner and, for each $i \in N$, we add an edge from the agent to his most preferred feasible continuation object in $U_i(q)$ (according to the linear representation of R_i we explained in Remark 1).

Implementation step. A trading cycle is a directed cycle in graph G(q). Given the finite number of nodes, at least one trading cycle exists. We assign to each agent *i* in a trading cycle the object that he pointed to, and denote the object assigned to him in this step by $a_i(q)$; we denote the corresponding set of objects assigned through trading cycles by

⁵See Konishi et al. (2001, Example 2.3) for details.

7

⁶The preference domain that Sikdar et al. (2017, 2019) consider is larger than ours.

A(q). If agent $i \in N$ was part of a trading cycle, then his partial allotment $x_i(q) = x_i(q-1) \cup \{a_i(q)\}$; otherwise, $x_i(q) = x_i(q-1)$.

Removal step. We remove all agents that have received a (complete) allotment and denote the set of remaining agents by N(q + 1). Next, we remove all objects that were assigned through trading cycles from set U(q) and set $U(q + 1) \coloneqq U(q) \setminus A(q)$. For each agent $i \in N(q)$, we derive the set of *feasible continuation objects* $U_i(q + 1)$ by removing all objects in U(q + 1) that are of a type that is already present in agent *i*'s partial allotment $x_i(q)$. Go to step q + 1.

Output. The MTTC algorithm terminates when all objects in *O* are assigned (it takes at most $n \cdot m$ steps). Assume that the final step is step q^* . Then, the final allocation is $x(q^*) = \{x_1(q^*), ..., x_n(q^*)\}.$

The *MTTC mechanism*, f^{MTTC} , assigns to each problem $R \in \mathcal{R}_l^N$ the allocation $x(q^*)$ obtained by the MTTC algorithm.

Sikdar et al. (2017, Theorem 1) proved that f^{MTTC} is strictly core stable but not strategyproof on the domain of lexicographically separable preferences \mathcal{R}_l^N . We illustrate the MTTC mechanism with Example 3 in Appendix A in the online supplementary materials.

2.3 | Preference revelation games

We now formulate a natural preference revelation game for the domain of lexicographically separable preferences (\mathcal{R}_l) . Given a multiple-type housing market problem represented by $R \in \mathcal{R}_l^N$ and a mechanism $f: \mathcal{R}_l^N \to X$, the *preference revelation game* induced by f is the *strategic game* $\Gamma_f(R) = (\mathcal{R}_l^N, f, R)$, where \mathcal{R}_l is each agent's *strategy space*, f is the *outcome function*, and each agent i evaluates outcomes with R_i .

Finding suitable solutions of strategic games is an important task in game theory. The most studied solution concept is the Nash equilibrium. However, Nash equilibria may not be Pareto efficient (e.g., in prisoner's dilemma or tragedy of the commons situations). To reestablish Pareto efficiency, Aumann (1959) introduced strong Nash equilibria, a strengthening of Nash equilibria that requires an equilibrium strategy profile to be robust against coalitional deviations (see footnote 7). By requiring Pareto efficiency for each coalition, Dubey (1986) introduced an even stronger refinement of the set of Nash equilibria: strict strong Nash equilibria.

Definition 8 (Nash/strict strong Nash equilibria). Let $R \in \mathcal{R}_l^N$ be a multiple-type housing market problem and consider the preference revelation game $\Gamma_f(R)$.

A strategy profile $R^* \in \mathcal{R}_l^N$ is a *Nash equilibrium* of $\Gamma_f(R)$ if for each agent $i \in N$ and each strategy $R'_i \in \mathcal{R}_l, f_i(R^*) = f_i(R^*_i, R^*_{-i})R_i f_i(R'_i, R^*_{-i})$. We denote the set of *Nash* equilibria by $\operatorname{Nash}(\Gamma_f(R))$ and the set of *Nash equilibrium outcomes* by $f(\operatorname{Nash}(\Gamma_f(R)))$. A strategy profile $R^* \in \mathcal{R}_l^N$ is a strict strong Nash equilibrium⁷ of $\Gamma_f(R)$ if for each coalition $S \subseteq N$ and each strategy list $R'_S \in \mathcal{R}_l^S$, [for each agent $i \in S$, $f_i(R'_S, R^*_{-S})R_i f_i(R^*_S, R^*_{-S})$] implies [for each agent $i \in S$, $f_i(R'_S, R^*_{-S}) = f_i(R^*_S, R^*_{-S})$]. We denote the set of strict strong Nash equilibria by $SNash(\Gamma_f(R))$ and the set of strict strong Nash equilibrium outcomes by $f(sNash(\Gamma_f(R)))$.

Note that $\operatorname{sNash}(\Gamma_f(R)) \subseteq \operatorname{Nash}(\Gamma_f(R)) \subseteq \mathcal{R}_l^N$.

Given a preference revelation game $\Gamma_f(R)$, we say that agent *i* plays a *truth-telling strategy* if he truthfully reports his preferences R_i . If all agents play truth-telling strategies, then $R = (R_i)_{i \in N}$ is a *truth-telling strategy profile* at $\Gamma_f(R)$. Note that if *f* is strategy-proof, then truthtelling is a weakly dominant strategy for each agent and the truth-telling strategy profile is a weakly dominant strategy Nash equilibrium.

3 | RESULTS

3.1 | Motivating examples

As mentioned in the introduction, for Shapley–Scarf housing markets with strict preferences, the unique strict core allocation can be obtained by a unique individually rational, Pareto efficient, and strategy-proof mechanism (Ma, 1994), the TTC mechanism. Later, Sönmez (1999) considered a generalization of Shapley and Scarfs (1974) housing market problems, *generalized indivisible goods allocation problems* (see Appendix C in the online supplementary materials), and showed that, whenever the preference domain satisfies a certain condition of richness and if there exists a mechanism satisfying individual rationality, Pareto efficiency, and strategy-proofness, then for any problem having a non-empty strict core, the strict core must be essentially single-valued⁸ and the mechanism must choose a strict core allocation. Takamiya (2003) showed the following converse result: whenever the preference domain satisfies a certain condition of richness and if the strict core solution is essentially single-valued, then any selection from the strict core solution is strategy-proof.

However, for multiple-type housing market problems, these results do not hold anymore: Konishi et al. (2001) (Sikdar et al., 2017, respectively) showed that on the domain of separable preferences (lexicographically separable preferences, respectively), no mechanism satisfies individual rationality, Pareto efficiency, and strategy-proofness. Note that neither the domain of separable preferences nor the domain of lexicographically separable preferences satisfies the domain richness condition of Sönmez (1999) (see Appendix C in the online supplementary materials).

The following example shows that on the one hand an individually rational and Pareto efficient mechanism can pick an allocation at which no agent has an incentive to misrepresent his preferences, while on the other hand the strict core may be multi-valued (without being essentially single-valued).

⁷The set of strict strong Nash equilibria is a refinement of the set of strong Nash equilibria: a strategy profile $R^* \in \mathcal{R}_l^N$ is a strong Nash equilibrium of $\Gamma_f(R)$ if for each coalition $S \subseteq N$ and each strategy list $R'_S \in \mathcal{R}_l^S$, [for each agent $i \in S, f_i(R'_S, R^*_{-S})R_if_i(R^*_S, R^*_{-S})$] implies [for some agent $j \in S, f_j(R'_S, R^*_{-S}) = f_j(R^*_S, R^*_{-S})$]. For a discussion of existence of strict strong Nash equilibria we refer to Remark 2.

⁸The strict core is essentially single-valued if each agent is indifferent between any two strict core allocations.

10

FENG AND KLAUS

Example 1 (Non-manipulability and a multi-valued strict core). Consider $R \in \mathcal{R}_l^N$ with $N = \{1, 2, 3\}, T = \{H(ouse), C(ar)\}, O = \{H_1, H_2, H_3, C_1, C_2, C_3\}$, each agent *i*'s endowment (H_i, C_i) , and

 $R_1 : H_2, H_1, H_3, C_3, C_1, C_2,$ $R_2 : H_3, H_2, H_1, C_1, C_2, C_3,$ $R_3 : H_2, H_3, H_1, C_1, C_3, C_2.$

Applying the MTTC algorithm to *R*, at step 1, the trading cycle $2 \rightarrow H_3 \rightarrow 3 \rightarrow H_2 \rightarrow 2$ forms; the trading cycle at step 2 is $1 \rightarrow H_1 \rightarrow 1$; the trading cycle at step 3 is $1 \rightarrow C_3 \rightarrow 3 \rightarrow C_1 \rightarrow 1$; and at step 4, we have $2 \rightarrow C_2 \rightarrow 2$. The final outcome is the strict core allocation $x = ((H_1, C_3), (H_3, C_2), (H_2, C_1))$.

Note that at problem *R*, no agent has an incentive to misrepresent his preferences: agent 3 has no incentive to misreport his preferences because he receives his best allotment. Agent 1 cannot obtain his best house H_2 by misreporting his preferences (it is traded in step 1 between agents 2 and 3). Given that, he receives the best possible allotment and has no incentive to misreport his preferences. Finally, agent 2 already obtains his best house, and if he tries to obtain his best car by misreporting his preferences he cannot obtain his best house; thus, he has no incentive to misreport his preferences. Finally, the strict core is not unique: $((H_1, C_3), (H_3, C_1), (H_2, C_2))$ is also a strict core allocation.

Recall that for multiple-type housing market problems with lexicographically separable preferences, no mechanism satisfies individual rationality, Pareto efficiency, and strategy-proofness (Sikdar et al., 2017). Hence, strict core stability and strategy-proofness are also not compatible. Thus, in our context, strategy-proofness, or truth-telling being a weakly dominant strategy Nash equilibrium in the corresponding preference revelation game, is a very strong requirement. Therefore, we next consider implementation through a different equilibrium concept: strict strong Nash equilibrium.

For generalized indivisible goods allocation problems, Takamiya (2009) studied the relationship between coalitional equilibria and the strict core. Takamiya's main result implies that for Shapley–Scarf housing market problems and for a preference revelation game induced by an individually rational and Pareto efficient mechanism f, the set of strict strong Nash equilibrium outcomes equals the strict core.

Corollary 1 (Takamiya, 2009). For each Shapley–Scarf housing market $R \in \mathcal{R}^N$ and each individually rational and Pareto efficient mechanism $f, f(\operatorname{sNash}(\Gamma_f(R))) = \operatorname{SC}(R)$.

The following example shows that Corollary 1 does not extend to multiple-type housing markets with lexicographically separable preferences.

Example 2 (Corollary 1 does not extend to R_l^N). Consider $R \in \mathcal{R}_l^N$, $N = \{1, 2, 3\}$, $T = \{H(ouse), C(ar)\}$, $O = \{H_1, H_2, H_3, C_1, C_2, C_3\}$, each agent *i*'s endowment (H_i, C_i) , and

 $R_1 : H_2, H_1, H_3, C_3, C_2, C_1,$ $R_2 : H_1, H_2, H_3, C_1, C_2, C_3,$ $R_3 : H_1, H_3, H_2, C_1, C_3, C_2.$

Applying the MTTC algorithm to *R*, at step 1, the trading cycle $1 \rightarrow H_2 \rightarrow 2 \rightarrow H_1 \rightarrow 1$ forms; the trading cycle at step 2 is $3 \rightarrow H_3 \rightarrow 3$; the trading cycle at step 3 is $1 \rightarrow C_3 \rightarrow C_3$ $3 \rightarrow C_1 \rightarrow 1$; and at step 4 we have $2 \rightarrow C_2 \rightarrow 2$. The final outcome is the strict core allocation $x = ((H_2, C_3), (H_1, C_2), (H_3, C_1))$. There is another strict core allocation x' = $((H_2, C_2), (H_1, C_1), (H_3, C_3))$. In Appendix A in the online supplementary materials we show that $\{x\} = f^{\text{MTTC}}(\text{sNash}(\Gamma_{f^{\text{MTTC}}}(R))) \subseteq \text{SC}(R) = \{x, x'\}.$

Based on Corollary 1 and Example 2, one could now conjecture that for each multiple-type housing market $R \in \mathcal{R}_l^N$ and each individually rational and Pareto efficient mechanism f, we have $f(\operatorname{sNash}(\Gamma_f(R))) \subseteq \operatorname{SC}(R)$. That conjecture is almost correct; however, we need to strengthen individual rationality and Pareto efficiency to strict core stability (see Example 6 in Appendix A in the online supplementary materials).

Main results 3.2

We show that for lexicographically separable preferences, if a mechanism is strictly core stable, then any strict strong Nash equilibrium of the corresponding preference revelation game will induce a strict core allocation. However, for some lexicographically separable multiple-type housing markets, there exist strict core allocations that cannot be implemented in strict strong Nash equilibrium.

Theorem 1. Let f be a strictly core stable mechanism on \mathcal{R}_l^N . Then, for each problem $R \in \mathcal{R}_l^N$ and the corresponding preference revelation game $\Gamma_f(R) = (\mathcal{R}_l^N, f, R)$, the set of strict strong Nash equilibrium outcomes is a subset of the strict core, that is, $f(sNash(\Gamma_f(R)))$ \subseteq SC(R). Furthermore, there exist problems $R \in \mathcal{R}_{l}^{N}$ such that $f(sNash(\Gamma_{f}(R))) \subseteq SC(R)$.

We would like to emphasize that the strict core stability of f is key for this result. Clearly, if for some preference profiles the strict core is empty, then a strictly core stable mechanism fcannot exist. Thus, in this first result, we restrict the preference domain to \mathcal{R}_{l}^{N} with the intent to generalize Theorem 1 later on.

Proof. Let f be a strictly core stable mechanism on \mathcal{R}_l^N . First, let $R \in \mathcal{R}_l^N$ and assume, by contradiction, that $f(\mathrm{sNash}(\Gamma_f(R))) \nsubseteq \mathrm{SC}(R)$. Let $R' \in \mathcal{R}_l^N$ be such that $R' \in \mathrm{sNash}(\Gamma_f(R))$ and $f(R') = x \notin SC(R)$. Hence, x can be weakly blocked by a coalition S and there exists an allocation y such that (1) for each $i \in S$ and each $t \in T$, $y_i^t \in \{o_j^t\}_{i \in S}$, and (2') for each $i \in S$, $y_i R_i x_i$, and for some $j \in S$, $y_j P_j x_j$.

Now we consider the profile $(\hat{R}_S, R'_{-S}) \in \mathcal{R}_l^N$ such that each agent $i \in S$ ranks allotment y_i as his best allotment; for each $i \in S$, we then have that $\hat{R}_i : y_i, ...,$ that is, each agent i, for each object type t, ranks y_i^t as best type-t object. We want to show that coalition S has an incentive to deviate from R'_S to \hat{R}_S . To this end, we first prove the following claim.

Claim 1. For each $i \in S$, we have $f_i(\hat{R}_S, R'_{-S}) = y_i$. Let $z = f(\hat{R}_S, R'_{-S})$. Suppose that for some agent $j \in S$, $z_j \neq y_j$. We show that z is not a strict core allocation at (\hat{R}_S, R'_{-S}) , that is, $z \notin SC(\hat{R}_S, R'_{-S})$.

WILEY

At (\hat{R}_S, R'_{-S}) , for each agent $i \in S$, $y_i \hat{R}_i z_i$ because y_i is his best allotment. Since $z_j \neq y_j$, $y_j \hat{P}_j z_j$. Therefore, at (\hat{R}_S, R'_{-S}) , allocation z can also be weakly blocked by coalition S via allocation y. Thus, $f(\hat{R}_S, R'_{-S}) \notin SC(\hat{R}_S, R'_{-S})$, which contradicts that f is strictly core stable.

Strictly speaking, by Claim 1, we now only know that $f(\hat{R}_S, R'_{-S}) = y'$ such that $y'_S = y_S$. However, since allotments to agents in $N \setminus S$ play no role in our proof, it is without loss of generality to assume that y' = y. Hence, when coalition S deviates from R'_{S} to \hat{R}_{S} , by Claim 1 and without loss of generality, $f(\hat{R}_{S}, R'_{-S}) = y$. Thus, since f(R') is weakly blocked by S via y, for each $i \in S$, $f_i(\hat{R}_S, R'_{-S})R_i f_i(R')$ and for $j \in S$,

 $f_j(\hat{R}_S, R'_{-S})P_j f_j(R')$; contradicting that R' is a strict strong Nash equilibrium.

Example 2 exhibits a problem $R \in \mathcal{R}_l^N$ such that $f(sNash(\Gamma_f(R))) \subsetneq SC(R)$ (recall that in Example 2 there is a unique strict strong Nash equilibrium outcome while multiple strict core allocations exist).

Remark 2 (Existence of strict strong Nash equilibria: an open problem). The existence of (strict) strong Nash equilibria has been proven for specific classes of games, such as congestion games (Holzman and Law-Yone, 1997), cost-sharing games (Epstein et al., 2009), and continuously convex games (Nessah and Tian, 2014). However, in general, (strict) strong Nash equilibria need not to exist.9

Question: Let f be a strictly core stable mechanism on \mathcal{R}_l^N . For each problem $R \in \mathcal{R}_l^N$, do we have $f(\operatorname{sNash}(\Gamma_f(R))) \neq \emptyset$?

For Shapley-Scarf housing markets and the TTC mechanism, truth-telling is a strict strong Nash equilibrium. Thus, for higher-dimensional multiple-type housing markets, one could conjecture that for f^{MTTC} , MTTC allocations can always be implemented in strict strong Nash equilibrium. Example 4 (Appendix A in the online supplementary materials) shows that the MTTC allocation cannot always be implemented truthfully in strict strong Nash equilibrium: an implementation of the MTTC allocation in strict strong Nash equilibrium might require some agents to (possibly mutually) change their object type sequences. We neither found a systematic way for agents to change their object type sequences to show existence of strict strong Nash equilibria, nor did we manage to construct a counter-example.

A more general result 3.3

Note that the proof of Theorem 1 did not use many properties of the lexicographically separable preference domain. It turns out that our result can easily be extended to other preference domains. Consider a subdomain of preferences $\hat{\mathcal{R}} \subseteq \mathcal{R}$ that satisfies the following two assumptions.

⁹Hoefer and Skopalik (2013) pointed out the following technical difficulty of finding strong Nash equilibria: "a strong Nash equilibrium must be the optimal solution of multiple non-convex optimization problems."

WILEY

Assumption 1 (Strict core existence and minimal preference domain richness). Preference domain $\hat{\mathcal{R}} \subseteq \mathcal{R}$ satisfies:

- (a) strict core existence if for each problem $R \in \hat{\mathcal{R}}^N$, SC(R) $\neq \emptyset$; and
- (**b**) *minimal preference domain richness* if for each allocation $x \in X$, each agent *i* can position x_i as his best allotment, that is, for each $x \in X$, there exists a profile $\hat{R} \in \hat{\mathcal{R}}^N$ such that for each $i \in N$, $\hat{R}_i : x_i$,

Assumption 1 is simple and reasonable. Assumption 1(a) allows us to focus on the solution of the strict core, and for that the strict core should always be non-empty. Assumption 1(b) is a very weak preference domain richness condition that is different from the one used by Sönmez (1999, Assumption B) and weaker than the one imposed by Takamiya (2009, Condition A). We discuss the preference domain richness conditions of Sönmez (1999) and Takamiya (2009) in Appendix C in the online supplementary materials.

Remark 3 (Preference domains satisfying Assumption 1). The domains of weak and strict preferences for Shapley–Scarf housing markets and the lexicographically separable preference domain for multiple-type housing markets all satisfy Assumption 1. There are various larger lexicographic domains, for example, those of Monte and Tumennasan (2015; generalized lexicographical preferences) and Sikdar et al. (2017; lexicographical preferences), that satisfy Assumption 1. Hence, our Theorem 1 applies to these settings as well (see the following Theorem 2).

We now show that Theorem 1 can be extended to any preference domain $\hat{\mathcal{R}} \subseteq \mathcal{R}$ satisfying Assumption 1.

Theorem 2. Let $\hat{\mathcal{R}}$ satisfy Assumption 1 and let f be a strictly core stable mechanism on $\hat{\mathcal{R}}^N$. Then, for each problem $R \in \hat{\mathcal{R}}^N$ and the corresponding preference revelation game $\Gamma_f(R) = (\hat{\mathcal{R}}^N, f, R)$, the set of strict strong Nash equilibrium outcomes is a subset of the strict core, that is, $f(\operatorname{sNash}(\Gamma_f(R))) \subseteq \operatorname{SC}(R)$. Furthermore, there exist problems $R \in \hat{\mathcal{R}}^N$ such that $f(\operatorname{sNash}(\Gamma_f(R))) \subseteq \operatorname{SC}(R)$.

Proof. The proof is the same as that of Theorem 1 since in that proof the only properties of the preference domain that were (implicitly) used were strict core existence and minimal domain richness.

We discuss the role that assumptions in Theorems 1 and 2 play in Appendix B in the online supplementary materials.

4 | CONCLUSION

We consider multiple-type housing market problems when agents have lexicographically separable preferences \mathcal{R}_l ; or alternatively, preferences are drawn from a preference domain $\hat{\mathcal{R}}$ that guarantees strict core existence and that satisfies a minimal preference domain richness condition (see Assumption 1). We show that if a mechanism is strictly core stable, then any strict strong Nash equilibrium outcome of its corresponding preference revelation game is a strict core allocation (Theorems 1 and 2). The converse statement is not true, that is, there exist problems with strict core allocations that cannot be implemented in strict strong Nash equilibrium (Example 2). We also demonstrated the necessity of two crucial assumptions (strict core non-emptiness and strict core stability of mechanisms) in our results (Examples 5 and 6 in the online supplementary materials).

Comparing our results to Takamiya's result for Shapley–Scarf housing markets, Corollary 1 (see Appendix C in the online supplementary materials for the generalized individual goods allocation model considered in Takamiya, 2009), our results (Theorems 1 and 2) have two differences with respect to his main result.

First, we show that not all strict core allocations may be implementable through strict strong Nash equilibria of the preference revelation game, while Takamiya (2009) showed full implementation for Shapley–Scarf housing markets. The main reason for our partial implementation versus his full implementation result is that our preference domains are less rich than the ones he considers. Neither separable nor lexicographically separable preferences satisfy Takamiya's preference domain richness condition (Takamiya, 2009, Condition A, see Appendix C in the online supplementary materials). For example, for multiple-type housing market problems with lexicographically separable preferences, no agent can protect an allotment by positioning it as his first best and his endowment as his second best allotment (an argument that is crucial in Takamiya's proof).

Second, we require strict core stability for our mechanisms while Takamiya (2009) only required individual rationality and Pareto efficiency. In Takamiya's model, each agent only demands one object. Thus, each agent will trade within only one coalition. Therefore, once the induced allocation (the equilibrium outcome) is individually rational and Pareto efficient, no coalition can block it.¹⁰ However, the same is not true for multiple-type housing market problems because each agent may trade different objects with different coalitions. That is, a multiple-type housing market problem cannot be easily transformed into a coalition formation game.

ACKNOWLEDGMENTS

We gratefully acknowledge financial support from the Swiss National Science Foundation (SNFS), project 100018_192583. We thank an anonymous referee, Flip Klijn, and Seçkin Özbilen for their valuable comments. And given the occasion, thank you, William: you are an amazing mentor and colleague, and hopefully this paper reflects some of the skills you taught us over the years (even if after a careful reading, you'd probably still return this paper with many suggestions for improvement (2).

REFERENCES

14

- Aumann, R.J. (1959) Acceptable points in general cooperative n-person games. In: Tucker, A.W. & Luce, R.D. (Eds.) Contributions to the theory of games (AM-40), Volume IV, Princeton, NJ: Princeton University Press, pp. 287–324.
- Banerjee, S., Konishi, H. & Sönmez, T. (2001) Core in a simple coalition formation game. Social Choice and Welfare, 18(1), 135–153.
- Bogomolnaia, A. & Jackson, M.O. (2002) The stability of hedonic coalition structures. Games and Economic Behavior, 38(2), 201–230.
- Dubey, P. (1986) Inefficiency of Nash equilibria. Mathematics of Operations Research, 11(1), 1-8.

- Ehlers, L. & Klaus, B. (2003) Coalitional strategy-proof and resource-monotonic solutions for multiple assignment problems. *Social Choice and Welfare*, 21(2), 265–280.
- Epstein, A., Feldman, M. & Mansour, Y. (2009) Strong equilibrium in cost sharing connection games. Games and Economic Behavior, 67(1), 51–68.
- Gale, D. & Shapley, L.S. (1962) College admissions and the stability of marriage. *American Mathematical Monthly*, 69, 9–15.
- Hoefer, M. & Skopalik, A. (2013) On the complexity of Pareto-optimal Nash and strong equilibria. Theory of Computing Systems, 53(3), 441–453.
- Holzman, R. & Law-Yone, N. (1997) Strong equilibrium in congestion games. *Games and Economic Behavior*, 21(1-2), 85–101.
- Huh, W.T., Liu, N. & Truong, V.-A. (2013) Multiresource allocation scheduling in dynamic environments. Manufacturing & Service Operations Management, 15(2), 280–291.
- Jackson, M.O. & Wolinsky, A. (1996) A strategic model of social and economic networks. Journal of Economic Theory, 71(1), 44–74.
- Klaus, B. (2008) The coordinatewise core for multiple-type housing markets is second-best incentive compatible. Journal of Mathematical Economics, 44(9-10), 919–924.
- Konishi, H., Quint, T. & Wako, J. (2001) On the Shapley–Scarf economy: the case of multiple types of indivisible goods. *Journal of Mathematical Economics*, 35(1), 1–15.
- Ma, J. (1994) Strategy-proofness and the strict core in a market with indivisibilities. International Journal of Game Theory, 23(1), 75–83.
- Mackin, E. & Xia, L. (2016) Allocating indivisible items in categorized domains. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, 359–365.
- Manjunath, V. & Westkamp, A. (2021) Strategy-proof exchange under trichotomous preferences. Journal of Economic Theory, 193, 105197.
- Monte, D. & Tumennasan, N. (2015) Centralized allocation in multiple markets. Journal of Mathematical Economics, 61, 74–85.
- Moulin, H. (1995) Cooperative microeconomics: a game-theoretic introduction. Princeton, NJ: Princeton University Press.
- Nessah, R. & Tian, G. (2014) On the existence of strong Nash equilibria. Journal of Mathematical Analysis and Applications, 414(2), 871–885.
- Pápai, S. (2001) Strategyproof and nonbossy multiple assignments. *Journal of Public Economic Theory*, 3(3), 257–271.
- Pápai, S. (2007) Exchange in a general market with indivisible goods. *Journal of Economic Theory*, 132(1), 208–235.
- Quint, T. & Wako, J. (2004) On houseswapping, the strict core, segmentation, and linear programming. Mathematics of Operations Research, 29(4), 861–877.
- 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.
- Shapley, L.S. & Scarf, H. (1974) On cores and indivisibility. Journal of Mathematical Economics, 1(1), 23-37.
- Sikdar, S., Adalı, S. & Xia, L. (2017) Mechanism design for multi-type housing markets. In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, pp. 684–690.
- Sikdar, S., Adalı, S. & Xia, L. (2019) Mechanism design for multi-type housing markets with acceptable bundles. In: Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence, 2165–2172.
- Sönmez, T. (1999) Strategy-proofness and essentially single-valued cores. Econometrica, 67(3), 677-689.
- Takamiya, K. (2003) On strategy-proofness and essentially single-valued cores: A converse result. *Social Choice and Welfare*, 20(1), 77–83.
- Takamiya, K. (2009) Preference revelation games and strong cores of allocation problems with indivisibilities. *Journal of Mathematical Economics*, 45(1-2), 199–204.
- Wako, J. (2005) Coalition-proof Nash allocation in a barter game with multiple indivisible goods. *Mathematical Social Sciences*, 49(2), 179–199.

JJET -WILEY

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Feng, D. & Klaus, B. (2021). Preference revelation games and strict cores of multiple-type housing market problems. *International Journal of Economic Theory*, 1–16. https://doi.org/10.1111/ijet.12321