Economic Theory
https://doi.org/10.1007/s00199-022-01483-9

RESEARCH ARTICLE

Constrained-optimal tradewise-stable outcomes in the one-sided assignment game: a solution concept weaker than the core

David Pérez-Castrillo[1] · Marilda Sotomayor[2,3]

Received: 21 December 2021 / Accepted: 27 December 2022

Abstract
In the one-sided assignment game, any two agents can form a trade; they can decide to form a partnership and agree on how to share the surplus created. Contrary to the two-sided assignment game, stable outcomes often fail to exist in the one-sided assignment game. Hence the core, which coincides with the set of stable payoffs, may be empty. We introduce the idea of tradewise-stable (t-stable) outcomes: they are individually rational outcomes where all trades are stable; that is, no matched agent can form a blocking pair with any other agent, neither matched nor unmatched. We propose the set of constrained-optimal (optimal) t-stable outcomes, the set of the maximal elements of the set of t-stable outcomes, as a natural solution concept for this game. We prove

We thank Federico Echenique, Jean-Jacques Herings, the associate editor, and two reviewers for their constructive comments. We also thank the suggestions received at SING 15 in Turku, the 30th International Conference on Game Theory at Stony Brook, the ASSET meeting in Athens, the 6th World Congress of the Game Theory Society, and seminars at Aix-Marseille School of Economics, Paris School of Economics, Shanghai University of Finance and Economics, and Universidad Autónoma de Barcelona. Marilda Sotomayor acknowledges financial support from CNPq-Brazil. David Pérez-Castrillo acknowledges financial support from the Ministerio de Ciencia, Innovació n y Universidades and Feder (PGC2018-094348-B-I00), Generalitat de Catalunya (2017SGR-711), ICREA under the ICREA Academia programme, and Severo Ochoa Programme (CEX2019-000915-S). A previous version of this paper has circulated under the title "Conflict-free and Pareto-optimal allocations in the one-sided assignment game: A solution concept weaker than the core".

✉ David Pérez-Castrillo
david.perez@uab.es

Marilda Sotomayor
marildas@usp.br

1 Department of Economía e Hist. Económica, Edificio B, Barcelona School of Economics, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

2 Graduate School of Economics, Getulio Vargas Foundation-RJ, Praia de Botafogo, 190, Rio de Janeiro, RJ 22250-900, Brazil

3 Dep de Economia, Universidade de São Paulo-SP, Av. Prof. Luciano Gualberto, 908, Cidade Universitária, São Paulo, SP 5508-900, Brazil

Published online: 06 January 2023

that this set is non-empty, it coincides with the set of stable outcomes when the core is non-empty, and it satisfies similar properties to the set of stable outcomes even when the core is empty. We propose a partnership formation process that starts with the outcome where every player stands alone, goes through steps where the set of active players expands, always forming t-stable outcomes, and ends in an (in any) optimal t-stable outcome. Finally, we also use the new concept to establish conditions under which the core is non-empty.

Keywords Matching · Assignment game · Stability · Core · Trade · Tradewise-stable

JEL Classification C78 · D78

1 Introduction

In the one-sided assignment game of our title, there is a finite set of players, which we may think of being a set of innovative firms, a group of tenants, or an industry. The main activity of these players is to form two-player partnerships. For each partnership $\{i, j\}$, there is a non-negative number a_{ij}, available to the partners, which can be split between them the way both agree. The interpretation is that if a partnership $\{i, j\}$ is formed, then the partners accomplish a joint activity that generates an income a_{ij}. A partnership plus a division of the surplus created into two non-negative payoffs will be named a *trade*. If an agent does not form any partnership, he gets a zero payoff, and we say he is *inactive*. The set of all trades and the inactive players with their zero payoffs is a feasible outcome (an *outcome*, for short). Thus, an outcome (x, u) for this model involves both a one-to-one matching x (that is, a partition of the population in pairs of agents and singletons) and a vector of non-negative payoffs u (which reflects the sharing of the joint surplus for any two-agent partnership).

The theory developed here presupposes an environment where a one-sided assignment game takes place. The traditional assumptions are that all players are rational, present simultaneously in the market, freely communicate with each other, and make offers, counteroffers, and binding agreements. Moreover, preferences over outcomes, and the rules that govern any pairwise interactions, are common knowledge. To these assumptions, we add that all players have an "optimal cooperative behavior." That is, *players only engage in cooperation that is optimal for them*, in the sense that if they reach agreements, they should be sure that they cannot obtain more favorable options elsewhere. This assumption is plausible, given that the players make rational decisions and can freely communicate with each other. The questions that naturally emerge are: (a) What partnerships should form? And (b) how should the gains be divided between the partners in a trade? In other words, *what trades will be accomplished?*

The key concept to address the previous questions is that of a *stable trade*. A trade is stable if none of the agents involved is part of a blocking pair. As customary, we say that the pair of agents $\{i, j\}$ blocks the outcome (x, u) if the sum of their payoffs under u is less than their joint surplus a_{ij}.

The classic solution concept for our environment is *stability*. An outcome is stable if it does not have any blocking pair. Thus, at a stable outcome, all trades are stable, and two inactive players cannot block the outcome either.

Concerning this solution concept, we prove that *the set of stable payoffs coincide with the core* (Proposition 1) and may be empty (Example 1).[1] Furthermore, *any optimal matching*[2] *is compatible with any stable payoff*, and every matching in a stable outcome is optimal (Sotomayor 2005a, 2009a; Talman and Yang 2011). Consequently, *an inactive agent at a stable outcome has a zero payoff at every stable outcome*.

The outcomes where all trades are stable play a crucial role in defining our solution concept for the one-sided assignment game. We call them *tradewise-stable outcomes* (*t-stable outcomes* for short). Then, at a t-stable outcome, no matched agent can form a blocking pair with any other agent, neither matched nor unmatched. In particular, in a t-stable outcome, the set of active players, those who are matched, is a stable set. Therefore, the t-stability provides a kind of "internal stability." However, it does not always yield "external stability," given that a pair of inactive agents could have an incentive to deviate. Since the outcome where all players are inactive is t-stable, the set of t-stable outcomes is non-empty.

We show (Proposition 9) that any t-stable outcome has a coalitional structure (there may be more than one) that allows us to conjecture an acceptable process for the formation of such an outcome. According to this process, the given t-stable outcome is reached after a finite number of steps, starting from the t-stable outcome where everyone stands alone. In each step, one or more trades occur simultaneously, and the agents involved in each trade split the gains obtained in their partnership. Moreover, only stable trades will be made, consistent with the assumption of optimal cooperative behavior; hence, the outcomes yielded are t-stable. For the same reason, once a trade is made at a given stage, the agents involved will keep their agreements at the subsequent stages. Therefore, the t-stable outcome reached at a given stage "extends" the t-stable outcomes formed in the previous stages. Consequently, only inactive agents at a given stage trade among them at a subsequent one to increase their payoffs. Any such sequence will be called a *t-stable sequence*.

Since only t-stable outcomes are feasible in our environment, we can expect that starting from the outcome where everyone is inactive, the agents' trades always generate a t-stable sequence. A key characteristic of a t-stable sequence is that it always converges to a term that cannot be extended to another t-stable outcome. This final term is reached in a finite number of steps, given that the number of agents is finite. The end of the t-stable sequence leads to one of the following two cases.

1st. Case. No pair of agents can block, so the outcome is stable.

2nd. Case. There is at least one pair of inactive agents with a positive surplus, so the final outcome is unstable. However, the final outcome is "as stable as possible," in the sense that any new partnership would violate the optimal cooperative behavior of at least one of the agents involved.

[1] Gale and Shapley (1962) show that stable matchings may also not exist in the one-sided discrete model, that is, when utility is not transferable. The existence problem for that model was called by these authors "the roommate problem."

[2] A matching is optimal if it maximizes the total payoff in the set of feasible matchings.

The set of the maximal t-stable sequences and the set of t-stable outcomes that cannot be extended to any t-stable outcome coincide. This coalitional structure of the final outcomes of the t-stable sequences creates a new and non-empty solution concept for the one-sided assignment game, which is weaker than the set of stable outcomes. We show that such a solution concept is the set of t-stable outcomes that cannot be Pareto improved by any t-stable outcome. We call it the set of *constrained-optimal tradewise-stable outcomes* (*optimal t-stable outcomes*, for short). Therefore, the question posed above can be answered as follows: the stable trades that will be accomplished in the environment considered here are those belonging to optimal t-stable outcomes.

Since a stable outcome cannot be extended to a t-stable outcome, every stable outcome is a limit point of some t-stable sequence, so it is optimal t-stable. Moreover, we prove that *when the set of stable outcomes is non-empty, it coincides with the set of optimal t-stable outcomes* (Theorem 2). On the other hand, the set of optimal t-stable outcomes is always non-empty. (This is the sense in which the optimal t-stability concept is weaker than the stability concept.) Then, we can check if the set of stable outcomes is empty or not by examining a single outcome: if the limit point of any t-stable is unstable, then the set of stable outcomes is empty; otherwise, this set is non-empty.

Thus, our solution concept configures as the most natural solution concept for the one-sided assignment game. It is always non-empty. It proposes the set of stable outcomes when this set is non-empty and, when the set of stable outcomes is empty, it recommends a non-empty set of payoffs, as stable as possible.

We prove additional appealing properties of the optimal t-stable outcomes. In particular, we show that the set of optimal t-stable outcomes always has a structure similar to that of the set of stable outcomes. Indeed, *every optimal t-stable outcome provides the maximum total surplus among all t-stable outcomes, and no other t-stable outcome can achieve this level of total surplus*. Thus, the matchings that are compatible with optimal t-stable outcomes are "quasi-optimal." Furthermore, they are always part of an optimal matching (they are optimal matchings only if the core is non-empty, in which case the sets of quasi-optimal and optimal matchings coincide).

Moreover, as the stable outcomes, *each quasi-optimal matching is compatible with any optimal t-stable payoff*. Therefore, the set of optimal t-stable outcomes is the Cartesian product of the set of quasi-optimal matchings and the set of optimal t-stable payoffs. Consequently, *if an agent is unmatched at some optimal t-stable outcome, then he gets a zero payoff at every optimal t-stable outcome*.

We also use the relationship between the core and the set of optimal t-stable outcomes to establish conditions under which the core is non-empty. We show that the core is non-empty if and only if every optimal t-stable payoff is Pareto-optimal feasible (Theorem 3). Finally, we show that (when the core is empty) the set of blocking pairs of any optimal t-stable outcomes is the same, and we use this result to provide further conditions for the non-emptiness of the core (Theorem 4).

Our paper is related to two-sided matching models pioneered by Gale and Shapley (1962) (the marriage and the college admission models) and Shapley and Shubik (1971) (the two-sided assignment game). They provide an excellent framework to study pairwise interactions in environments where the players belong to two disjoint sets and interact by pairs.

There is a vast literature on two-sided matching models. However, very few papers have studied the one-sided assignment game.[3] Necessary and sufficient conditions for the existence of the core using linear programming are obtained by Talman and Yang (2011). Eriksson and Karlander (2000) use graph theory to characterize the core, and Klaus and Nichifor (2010) provide some properties of this set when it is not empty. Chiappori et al. (2014) show that stable matchings exist when the economy is replicated an even number of times by "cloning" each individual. Finally, Andersson et al. (2014) propose a dynamic competitive adjustment process that either leads to a stable outcome or disproves the existence of stable outcomes.

The idea of the partnership formation process that leads to an optimal t-stable outcome has some similarities with the "paths to stability" proposed by Roth and Vande Vate (1990). They analyze the marriage market and show that starting from an arbitrary matching, there always exists a process of myopic blocking pairs that leads to a stable matching. Klaus and Payot (2015) show that the existence of such paths to stability is not guaranteed in the assignment game and provide a necessary and sufficient condition for the existence. In contrast to the paths to stability that have been considered in various contexts, each step in our partnership formation processes may involve more than one pair. But every single step in the process produces an outcome with nice properties, as it is a t-stable outcome.

Our paper is also related to the literature that looks for non-empty solution concepts similar to the core. For transferable utility games, Kóczy and Lauwers (2007) and Herings and Kóczy (2021) introduce the "minimal dominant set," based on a bargaining process where players can propose and counter-propose outcomes (payoff vectors augmented with coalition structures) satisfying certain conditions. Demuynck et al. (2019) introduce the "myopic stable set" for a general class of social environments, extending the idea of level-1 farsighted stability (Herings et al. 2009). [4] Both solution concepts coincide for transferable utility games (Herings and Kóczy 2021). They also coincide with the coalition structure core when this set is not empty and, as is the case for the set of optimal t-stable outcomes, may lead to inefficient coalition structures when the core is empty.

The concept of tradewise-stable outcome is an extension of the concept of "simple outcome," introduced in Sotomayor (1996) for the marriage model, and it is a translation of the concept of a "simple outcome," defined in Sotomayor (2005b), for the housing market of Shapley and Scarf (1974), with strict or non-strict preferences. Both papers provide a non-constructive and concise proof, which only uses elementary combinatorial arguments, of the existence of the core. Still in environments without transfers, the notion of a simple outcome was used in Sotomayor (1999) for a discrete many-to-many matching model with substitutable and not-necessarily strict preferences, in Sotomayor (2004), where an implementation mechanism for the discrete many-to-many matching model is provided, in Sotomayor (2011) to characterize the set of Pareto-stable matchings in the marriage market and (if the set of stable match-

[3] It is called "the partnership formation problem" in Talman and Yang (2011) and Andersson et al. (2014), "the TU roommate game" in Eriksson and Karlander (2000), and simply "the roommate problem" in Chiappori et al. (2014).

[4] See also Herings et al. (2019) and Luo et al. (2021), who study the stability of networks with myopic and farsighted individuals.

ings is not empty) in the discrete roommate model, and in Wu and Roth (2018) for the college admission model. An adaptation of the concept of simple matching was used in Sotomayor (2000) for a unified two-sided matching model, due to Eriksson and Karlander (2000), which includes the marriage and the assignment model, and in Sotomayor (2018) for the two-sided assignment game of Shapley and Shubik (1971). Finally, Sotomayor (2019) introduces the idea of t-stability for one of the sides of the market and provides a framework to treat conjointly stable and unstable allocation structures.

The rest of the paper is organized as follows. Section 2 introduces the framework and states some preliminary results for the core. Section 3 introduces the set of optimal t-stable outcomes after defining and discussing the t-stable outcomes. Section 4 provides the main properties of the set of optimal t-stable outcomes. In Sect. 5, we introduce and discuss a dynamic partnership formation process that ends with an optimal t-stable outcome. In Sect. 6, we use the link between the set of optimal t-stable outcomes and the set of stable outcomes to provide conditions for the non-emptiness of the core. The final remarks are given in Sect. 7. An Appendix includes several proofs and lemmas that will be used in the proofs of the results in the main text.

2 Framework and preliminaries

2.1 The framework

The description of the one-sided assignment game follows the one given in Roth and Sotomayor (1990) for the case with two sides, with the appropriate adaptations.

There is a finite set of players, $N = \{1, 2, \ldots, n\}$. Associated with each partnership $\{i, j\}$ there is a non-negative real number $a_{\{i,j\}}$ which will denote a_{ij}. The number a_{ij} represents the surplus that players i and j generate if they form a partnership.

We can represent the environment as a game in coalitional function form (N, v) with side payments determined by (N, a). In this game, the worth $v(i, j)$[5] of a two-player coalition $\{i, j\}$ is given by a_{ij}. We define $v(i) \equiv a_{ii} \equiv 0$ for all $i \in N$. The worth of larger coalitions is entirely determined by the worth of the pairwise combinations that the coalition members can form. That is, $v(S) = max\{v(i_1, j_1) + v(i_2, j_2) + ... + v(i_k, j_k)\}$ for arbitrary coalitions S, where the maximum is taken over all sets $\{i_1, j_1\}, \ldots, \{i_k, j_k\}$ of two-player disjoint coalitions in S.[6]

Thus, the rules of the game are that any pair of agents $\{i, j\}$ can together obtain a_{ij}, and any larger coalition is valuable only insofar as it can organize itself into such pairs. The members of any coalition may divide their collective worth among themselves in any way they like.

We might think of the two-sided "Assignment Game" of Shapley and Shubik (1971) as a particular case of our model. There are two disjoint sets in the two-sided assignment game, P and Q and a pair of players can generate a surplus only if each belongs to a

[5] For notational convenience, we write $v(i, j)$ rather than $v(\{i, j\})$.

[6] k is an integer number that does not exceed the integer part of $|S|/2$.

different set. Thus, our model corresponds to an assignment game when $N = P \cup Q$, $P \cap Q = \varnothing$, and $v(S) = 0$ if S contains only agents of P or only agents of Q.

We represent the set of partnerships that are formed through a matching:

Definition 1 A *feasible matching* x is a partition of N, where the partition sets are either pairs $\{i, j\}$ or singletons $\{i\}$. If $\{i, j\} \in x$, we can write $x(i) = j$, and we refer to $x(i)$ as the partner of i at x. If $\{i\} \in x$ we can write $x(i) = i$ and we say that i is *unmatched* at x.

We use the notation $\sum_A$ to denote the sum over all elements of A. Let x be a feasible matching. If $R \subseteq N$, we denote $x(R) \equiv \{j : x(i) = j$ or some $i \in R\}$. If $x(R) = R$, we denote by $x|_R$ the partition of R where the partition sets belong to x. Therefore, $v(R) \geq \sum_{x|_R} a_{ij}$ for all feasible matchings x such that $x(R) = R$.

Definition 2 The feasible matching x is *optimal* if, for all feasible matching x', $\sum_x a_{ij} \geq \sum_{x'} a_{ij}$.

The set of optimal matchings is always non-empty since there is a finite number of matchings. Under Definition 2 and since $v(N) \geq \sum_x a_{ij}$ for all feasible matchings x, it follows that the matching x *is optimal if and only if* $\sum_x a_{ij} = v(N)$.

A vector of payoffs represents the players' benefit in the game:

Definition 3 The vector u, with $u \in \mathbb{R}^n$, is called the *payoff*. The payoff u is *pairwise-feasible* for (N, a) if there is a feasible matching x such that

$$u_i + u_j = a_{ij} \text{ if } x(i) = j \text{ and } u_i = 0 \text{ if } x(i) = i\ .$$

In this case, we say that (x, u) is a *pairwise-feasible outcome* and x is compatible with u.

That is, a pairwise-feasible outcome describes the set of all matched pairs, the sharing of the surplus generated by these pairs, and the unmatched agents with their zero payoffs.

When we describe the game in its coalitional function form, the corresponding concept is the feasibility of a payoff.

Definition 4 The payoff u is feasible for (N, a) if $\sum_N u_i \leq v(N)$.

The concept of feasibility is weaker than that of pairwise feasibility. In fact, the sum of payoffs of the members of any coalition in a pairwise-feasible outcome is always lower than or equal to the worth of the coalition. We state this property, which will be helpful later, in Remark 1.

Remark 1 Consider a coalition R. The definition of v implies that there is some feasible matching x such that $x(R) = R$ and $\sum_{x|_R} a_{ij} = v(R)$. Furthermore, $v(R) \geq \sum_{x'|_R} a_{ij}$ for all feasible matchings x' such that $x'(R) = R$. Then, it follows from Definition 3 that $\sum_R u_i \leq v(R)$ for all $R \subseteq N$ and for all pairwise-feasible outcomes (x, u) with $x(R) = R$. In particular, $\sum_N u_i \leq v(N)$, which implies that *every pairwise-feasible payoff is feasible.*

The natural solution concept is stability (the general definition of stability is given in Sotomayor 2009b). For the one-sided assignment game, stability is equivalent to pairwise stability.

Definition 5 The pairwise-feasible payoff u is *stable* if

(a) $u_i \geq 0$ for all $i \in N$ and
(b) $u_i + u_j \geq a_{ij}$ for all $\{i, j\} \subseteq N$.
If x is compatible with u we say that (x, u) is a stable outcome.

Condition (a) (individual rationality) means that in a stable situation, a player always has the option of remaining unmatched.[7] Condition (b) ensures the stability of the payoff distribution: If it is not satisfied for some agents i and j, then it would pay for them to break up their present partnership(s) and form a new one together, as this would give them each a higher payoff. In this case, we say that $\{i, j\}$ *blocks* u.

Finally, going back to the coalitional function form representation of the game, we define the core of (N, a), which we denote by C:

Definition 6 We say that $u \in C$ if $\sum_N u_i = v(N)$ and $\sum_S u_i \geq v(S)$ for all $S \subseteq N$.

The following example shows that the core of this model may be empty.

Example 1 Consider $N = \{1, 2, 3\}$ and $a_{ij} = 1$ for all $\{i, j\} \subset N$. For every feasible payoff u , there exist two players i and j such that $u_i + u_j < 1$. Hence, the core of this game is empty.

2.2 Preliminary results for the core

The concepts of stability of the payoffs and the core are equivalent in our environment, as established in the following proposition.

Proposition 1 *The set of stable payoffs coincides with the core of* (N, a).

Proof Suppose u is a stable payoff. Then, u is feasible, and so

$$\sum_N u_i \leq v(N) \tag{1}$$

according to Remark 1. Moreover, consider any coalition S and let y be a feasible matching such that $y(S) = S$ and $v(S) = \sum_y a_{ij}$. The stability of u implies that $u_i + u_{y(i)} \geq a_{iy(i)}$ for all $i \in S$, so

$$\sum_S u_i \geq v(S) \text{ for all coalition } S. \tag{2}$$

[7] We require individual rationality because our definition of a payoff is a vector $u \in \mathbb{R}^n$. Alternatively, we could have defined a payoff as a vector $u \in \mathbb{R}^n_+$. All our analysis would be the same if we would require from the start that a payoff is a non-negative vector, except that the individual rationality condition would not be necessary.

Under (1) and (2) it follows that $\sum_N u_i = v(N)$ and $\sum_S u_i \geq v(S)$ for all $S \subseteq N$, so u is in the core.

Now, suppose u is in the core. Definition 6 implies that $u_i + u_j \geq v(i, j) = a_{ij}$ for every coalition $\{i, j\}$ and $u_i \geq v(i) = 0$ for all $i \in N$, so u does not have any blocking pair and is individually rational. To see that u is pairwise-feasible, let x be a feasible matching such that $v(N) = \sum_x a_{ij}$. Use that $\sum_N u_i = v(N)$ and $u_i + u_{x(i)} \geq a_{ix(i)}$ for all $i \in N$, to get that $\sum_N u_i = \sum_x a_{ij} = \sum_{i<x(i)} a_{ix(i)} \leq \sum_{i<x(i)}(u_i + u_{x(i)}) \leq \sum_N u_i$, so the inequalities cannot be strict, which implies $u_i + u_{x(i)} = a_{ix(i)}$ for all $i \in N$ with $x(i) \neq i$. Since $u_i \geq 0$, it follows that $u_i = 0$ if $x(i) = i$. Hence, u is stable and the proof is complete. □

The following proposition, proven by Sotomayor (2005a, 2009a) and Talman and Yang (2011), makes clear why, similarly to the two-sided assignment game and in contrast to the discrete version (the roommate-problem), we can concentrate on the payoffs to the players rather than on the underlying matching. Indeed, it shows that the set of stable outcomes is the Cartesian product of the set of stable payoffs and the set of optimal matchings. We state this proposition without its proof (the proof of part (a) of Proposition 2 is basically the same as the second part of the proof of Proposition 1).

Proposition 2

(a) If x is an optimal matching then it is compatible with any stable payoff u.
(b) If (x, u) is a stable outcome, then x is an optimal matching.

A consequence of Proposition 2(a) is that, similarly to the two-sided assignment game, every inactive agent in a stable outcome has a zero payoff at any stable outcome in the one-sided assignment game. Corollary 1 states this result.

Corollary 1 *Let x be an optimal matching. If i is unmatched at x, then $u_i = 0$ for all stable payoffs u.*[8]

Proof Let $u \in C$. Under Proposition 2(a), u is compatible with x, so $u_i = 0$ by the pairwise-feasibility of u. □

We note that we can relate Corollary 1 to the so-called "rural hospitals theorem" that holds for the set of stable matchings in NTU two-sided matching (Roth 1986). The rural hospitals theorem states that if a hospital does not fill all its positions in one stable matching, then the set of doctors that are associated to this hospital under any stable matching is the same. Therefore, changing the stable matching rule used to allocate doctors to hospitals can not improve the final allocation of "rural hospitals," which are typically less preferred than those in urban areas by medical graduates.

To see the link between Corollary 1 and the rural hospitals theorem, we can reinterpret Corollary 1 as follows: If player i is unmatched at a stable outcome, then his payoff is zero at any stable outcome. Hence, a player who does not "fill his position" in a stable outcome does not improve his payoff if a different "stable rule" is followed.

[8] This result was proved in Demange and Gale (1985) for a two-sided matching market where the utilities are continuous, so it applies to the two-sided assignment game of Shapley and Shubik (1971).

3 Constrained-optimal tradewise-stable outcomes

In this section, we introduce the set of constrained-optimal tradewise-stable (optimal t-stable, for short) outcomes. We propose this set as a solution concept for the one-sided assignment game. We see it as the natural solution concept if one cares about payoffs that are "as stable as possible." Indeed, as we will see in Sect. 6, the optimal t-stable outcomes are the limit points of t-stable sequences, where each term extends the previous ones through the formation of stables trades. Moreover, we show that the set of optimal t-stable outcomes satisfies several desirable properties (proven in this and the following section). First, the set of optimal t-stable outcomes is always non-empty, and it coincides with the set of stable outcomes when the core is non-empty. Second, every optimal t-stable outcome is "internally stable," so no "active player" has an incentive to look for other partners inside or outside the set of active players. Third, each optimal t-stable outcome provides the maximum surplus out of the set of internally stable outcomes; and all the internally stable outcomes outside the set provide less surplus. Fourth, any internally stable outcome that is not optimal t-stable can be naturally extended to an optimal t-stable outcome. Fifth, all optimal t-stable outcomes are compatible with the same matchings. Finally, the previous properties of optimal t-stable outcomes replicate properties that are satisfied by the stable outcomes.

Before we define optimal t-stable outcomes, we introduce and discuss the concepts of a trade, a stable trade, and a tradewise-stable outcome.

A partnership $\{i, j\}$ in a matching x plus a division of the surplus a_{ij} created into two non-negative payoffs u_i and u_j (with $u_i + u_j = a_{ij}$) is named a *trade*, which we represent through $(\{i, j\}; (u_i, u_j))$. Then, we define a stable trade as follows:

Definition 7 The *trade* $(\{i, j\}; (u_i, u_j))$ in the outcome (x, u) is *stable* if neither i nor j is part of a blocking pair of u.

That is, trade $(\{i, j\}; (u_i, u_j))$ is stable in the outcome (x, u) if $u_i + u_k \geq a_{ik}$ and $u_k + u_j \geq a_{kj}$ for every $k \in N \backslash \{i, j\}$.

We note that can use the idea of stable trades to provide an equivalent definition of stable outcomes in our framework. We say that an agent is *inactive* if he does not form any partnership in x (hence, he gets a zero payoff). Then, it is immediate to check that Definition 8 below is equivalent to the definition of a stable outcome in Definition 5.

Definition 8 The pairwise-feasible outcome (x, u) is *stable* if it is individually rational and

(a) All trades $(\{i, j\}; (u_i, u_j))$ in (x, u) are stable, and
(b) Two inactive agents cannot block (x, u).

The definition of a stable coalition will also be useful in what follows.

Definition 9 Let (x, u) be a pairwise-feasible outcome. Let $R \subseteq N$. We say that R is a *stable coalition* for (x, u) if (a) $x(R) = R$ and (b) $u_i + u_j \geq a_{ij}$ for all $\{i, j\} \subseteq R$.

Remark 2 Notice that if R is stable for (x, u) it must be the case that $\sum_R u_i \geq v(R)$, according to Definition 9. On the other hand, $\sum_R u_i \leq v(R)$, as stated in Remark 1. Therefore, $\sum_R \boldsymbol{u}_i = \boldsymbol{v}(\boldsymbol{R})$. □

We now introduce and discuss the concept of a tradewise-stable outcome. Tradewise-stable outcomes satisfy properties similar to, but weaker than, stable outcomes.

Definition 10 The pairwise-feasible outcome (x, u) is *tradewise-stable* (*t-stable*) if it is individually rational and all trades $(\{i, j\}; (u_i, u_j))$ in (x, u) are stable.

We say that the matching x is t-stable if there is some payoff u such that the outcome (x, u) is t-stable. Similarly, the payoff u is a *t-stable payoff* if there is some matching x such that (x, u) is a t-stable outcome.

Every stable outcome is t-stable. However, t-stable outcomes are not necessarily stable since they do not necessarily satisfy condition (b) of Definition 8. In a t-stable outcome, there may be inactive agents who could block the outcome. For instance, the outcome where every player is unmatched and obtains a payoff of 0 is t-stable, but it is not a stable outcome in any game where at least one partnership creates a positive surplus. We denote by (x^0, u^0) this outcome, that is, $x^0(i) = i$ and $u_i^0 = 0$ for all $i \in N$. Moreover, this example allows us to state that *the set of t-stable outcomes is always non-empty*.

To discuss the difference between stable and t-stable outcomes, consider a pairwise-feasible and individually rational outcome (x, u). We denote by $T(x)$ the set "active players," that is, the set of all players who are matched at x, and by $U(x)$ the set of inactive players at x. That is,

$$T(x) \equiv \{j \in N : x(j) \neq j\} \text{ and } U(x) \equiv N \backslash T(x).$$

The outcome (x, u) is t-stable if and only if no player in $T(x)$ can form a blocking pair neither with another player in $T(x)$ (which implies that the coalition $T(x)$ is a stable coalition for (x, u)) nor with any player in $U(x)$. In this sense, we could say that a t-stable outcome is "internally stable." We refer to $T(x)$ as the stable active coalition for (x, u). However, to be stable, the outcome also needs to be "externally stable," in the sense that no pair of players in $U(x)$ can block the outcome either. A t-stable outcome might not be externally stable.[9]

In any t-stable outcome (x, u), the set of active players $T(x)$ organize themselves in the best possible way in the sense that the sum of their payoffs is the maximum that they can achieve. On the other hand, if (x, u) is t-stable but not stable, then the payoff of every inactive player is 0, but the worth of the coalition of inactive players $U(x)$ is positive. We state and prove these two results in Remarks 3 and 4.

Remark 3 If (u, x) is t-stable then the set $T(x)$ and all the subsets $R \subseteq T(x)$ such that $x(R) = R$ are stable coalitions for (u, x). Therefore, by Remark 2, $\sum_{T(x)} u_i = v(T(x))$, and for such coalitions R, we also have $\sum_R u_i = v(R)$.

[9] This property of external stability is not related to the definition of stability as in von Neumann and Morgenstern (1944). It is more in line with (although different from) the idea of external stability in d'Aspremont et al. (1983).

Remark 4 If the t-stable outcome (u, x) is not stable then $v(U(x)) > 0$. Indeed, let $\{j, k\}$ be a blocking pair. Then $\{j, k\} \subseteq U(x)$, so

$$0 = \sum_{U(x)} u_i = (u_j + u_k) + \sum_{U(x)\backslash\{j,k\}} u_i < a_{jk} + \sum_{U(x)\backslash\{j,k\}} a_{ii} \leq v(U(x)).$$

Therefore, $v(U(x)) > 0$.

We denote by S *the set of t-stable payoffs*:

$$S \equiv \{u \in \mathbb{R}^n\colon u \text{ is t-stable}\}.$$

Our solution concept, that is, the set of *constrained-optimal* tradewise-stable outcomes, is the set of the t-stable outcomes that are not dominated, via coalition N, by any other t-stable outcome. To introduce the set, let us first formally define the notion of Pareto optimality.

Definition 11 Let A be a set of payoffs. The payoff u is *Pareto-optimal* (PO) *in* A (or among all payoffs in A) if it belongs to A and there is no payoff w in A such that $w > u$.[10]

Similarly, we say that (x, u) is a PO outcome in A if u is PO in A and x is compatible with u.

If A is the set of individually rational and feasible payoffs and u is PO in A, then we refer to u as a PO feasible payoff. It follows from Definition 11 that an individually rational and feasible payoff u is PO feasible if and only if $\sum_N u_i = v(N)$. Thus, every stable payoff is PO feasible. However, the Pareto optimality of a payoff is not enough to guarantee its stability. For instance, in Example 1, the payoff $u = (1, 0, 0)$ is not in the core, but it is PO feasible since $\sum_N u_i = v(N)$.

The case in which the set A in Definition 11 corresponds to the set of t-stable payoffs, that is, $A = S$, is the main focus of interest of our theory.

Definition 12 The payoff u is a *constrained-optimal tradewise-stable (optimal t-stable*, for short) *payoff* if it is a t-stable payoff and it is PO in the set of t-stable payoffs.
The outcome (x, u) is an *optimal t-stable outcome* if (x, u) is a t-stable outcome and u is an optimal t-stable payoff.

We denote the *set of optimal t-stable payoffs* by S^*:

$$S^* \equiv \{u \in S : u \text{ is Pareto optimal in } S\}.$$

Theorem 1 uses Lemmas 1 and 2 in the Appendix to state that the set S^* is always non-empty. Moreover, S^* is a compact set.[11]

[10] Given two vectors $w, v \in \mathbb{R}^n$, we will denote $w > u$ if $w_j \geq u_j$ for all players $j \in N$ and $w_j > u_j$ for at least one player $j \in N$.

[11] Lemma 1 in the Appendix shows that the set S is also a non-empty and compact set of $\mathbb{R}^n$.

Theorem 1 *The set of optimal t-stable payoffs S^* is a non-empty and compact set of $\mathbb{R}^n$.*

Proof According to Lemma 1, S is compact and non-empty. Moreover, S is an ordered set by the partial order relation $\geq$ induced by $\mathbb{R}^n$. Then, Lemma 2 applies, and so S^*, the set of maximal elements of S, is a non-empty and compact set of $\mathbb{R}^n$. □

Theorem 1 implies that the set of optimal t-stable outcomes is non-empty for any one-sided assignment game. In Example 1, where the core is empty, the agents cannot form any partnership in a t-stable manner. The unique optimal t-stable outcome is that the three players remain unmatched and obtain a zero payoff. Example 2 provides a more illustrative game with an empty core.

Example 2 The set of players is $N = \{1, 2, 3, 4, 5, 6, 7\}$, and the surplus of the partnerships is $a_{12} = 3, a_{34} = 2, a_{13} = 1, a_{56} = a_{57} = a_{67} = 1$, and $a_{ij} = 0$ for the other partnerships. The core of this game is empty because for every pairwise-feasible payoff u, there exist two players $i, j \in \{5, 6, 7\}$ with $u_i + u_j < 1$. The set of optimal t-stable outcomes is the set of outcomes (x, u) that satisfy $x(1) = 2$, $x(3) = 4$, and $x(i) = i$ for $i = 5, 6, 7$, and the payoffs are non-negative numbers with $u_1 + u_2 = 3$, $u_3 + u_4 = 2$, $u_1 + u_3 \geq 1$, and $u_5 = u_6 = u_7 = 0$.

Finally, we introduce the idea of an extension of a t-stable outcome, which will be helpful in the next section. In words, a t-stable outcome (z, w) extends the t-stable outcome (x, u) if all the players in the stable active coalition of (x, u) keep their payoff, but some players who were unmatched in (x, u) obtain a positive payoff (hence, they are matched) in (z, w).

Definition 13 Let (x, u) be a t-stable outcome. We say that the t-stable outcome (z, w) *extends* (x, u) if $T(x) \subsetneq T(z)$, $w_j = u_j$ for all $j \in T(x)$, and $w_k > u_k$ for some $k \in T(z) \backslash T(x)$.

We notice that, given Definition 13, every optimal t-stable outcome is non-extendable.

4 Properties of the set of constrained-optimal tradewise-stable outcomes

We have shown in the previous section that the set of optimal t-stable payoffs is always non-empty. In this section, we provide relevant properties of this set. The first main property (Theorem 2) is that the set of optimal t-stable payoffs coincides with the core (that is, with the set of stable payoffs) when the core is non-empty. Additional results (Propositions 5, 6, and 7) highlight that the set of optimal t-stable outcomes satisfies similar properties as the set of core outcomes, even when the core is empty. The last results in this section lay the foundation for the dynamic process introduced in Sect. 5.

Theorem 2 requires two previous results. First, Proposition 3 provides a relevant relationship between t-stable matchings and optimal matchings. It states that a t-stable matching may not be optimal, but it is always part of an optimal matching. It uses Lemma 3 (see the Appendix), which shows that $v(T(x)) + v(U(x)) = v(N)$ for any t-stable outcome (x, u) .

Proposition 3 *Let (x, u) be a t-stable outcome. Then, the set of active partnerships of x is part of an optimal matching.*

Proof The proof is immediate after the fact that $v(T(x)) + v(U(x)) = v(N)$ because of Lemma 3, $N = T(x) \cup U(x)$, and $T(x) \cap U(x) = \varnothing$. □

It is worth noticing that given that the set $T(x)$ is part of an optimal matching for any t-stable outcome (x, u) (Proposition 3) and that any optimal matching is compatible with any stable payoff (Proposition 2(b)), then it follows that $\sum_{T(x)} u_i = \sum_{T(x)} w_i$ for all $w \in C$ when the core is not empty. Hence, the sum of the payoffs of the set of agents matched in a t-stable outcome is maintained in any stable outcome.

Second, Proposition 4 proves an important properties of the t-stable outcomes when the core is non-empty. It states that for any t-stable outcome which is not stable, it is always possible to construct a stable outcome that keeps the payoff of each matched player. We could say that a t-stable outcome is "a part" of a stable outcome.

Proposition 4 *Let (x, u) be a t-stable outcome which is not a stable outcome. Suppose the set of stable outcomes is non-empty. Then there exists a stable outcome (z, u^*) that extends (x, u).*

Proof See the Appendix. □

The set of optimal t-stable payoffs provides a solution for every game. Theorem 2 uses Proposition 4 to state that this set coincides with the core, which is equivalent to the set of stable payoffs, if and only if the core is not empty. We will extend Theorem 2 from payoffs to outcomes in Corollary 2 once we analyze the relationship between the set of optimal t-stable payoffs and the set of optimal t-stable matchings.

Theorem 2 *The core is non-empty if and only if $S^* = C$.*

Proof Suppose the core is non-empty. Let (x, u) be an optimal t-stable outcome. We now show that (x, u) is a stable outcome. Suppose, by way of contradiction, that (x, u) is unstable. Under Proposition 4, there is some stable outcome (z, u^*) which extends (x, u). Then, $u^*_j \geq u_j$ for all $j \in N$ and $u^*_j > u_j$ for at least one player j. But this contradicts the fact that (x, u) is Pareto optimal in S. Hence, (x, u) is a stable outcome. Similarly, let (x, u) be a stable outcome. Then, (x, u) is PO feasible. Since every t-stable outcome is feasible, it follows that there is no t-stable Pareto improvement of (x, u). Given that (x, u) is t-stable, it must be an optimal t-stable outcome. Hence, $S^* = C$.

The proof that $S^* = C$ implies that the core is not empty is immediate from the fact that $S^* \neq \varnothing$. □

We have shown that our solution concept suggests the same payoffs as the most popular solution concept, the core, whenever the core is not empty. We now provide several appealing properties that it satisfies not only when it coincides with the core but for every game.

Proposition 5 states that if a player is not matched in an optimal t-stable outcome, then his payoff is zero (and the payoff of his partner is the same) at any optimal t-stable outcome. Hence, this player would not improve his situation even if he were matched at some other optimal t-stable matching. The proposition uses a Decomposition Lemma (Lemma 4), stated and proved in the Appendix.

Proposition 5 *Let* (x, u) *and* (y, w) *be optimal t-stable outcomes. Let* $j^* \in T(x)\backslash T(y)$. *Then* $u_{j^*} = w_{j^*} = 0$ *and* $u_{x(j^*)} = w_{x(j^*)}$.

Proof See the Appendix. □

We note that, after Theorem 2, Proposition 5 applies in particular to the core. Hence, it extends Corollary 1 (which reflects an idea close to the rural hospitals theorem and was proved for the core) to the set of optimal t-stable outcomes.

Next, we analyze the players' payoffs in optimal t-stable outcomes. Optimal t-stable payoffs are, by definition, undominated in the set S. Proposition 6 shows that all optimal t-stable outcomes are equally efficient, in the sense that the players' total payoff is the same in every optimal t-stable payoff. Moreover, the players' total payoff in every element of S^* is higher than in every t-stable payoff vector not in S^*. The proposition uses Lemma 5 (see the Appendix), which states that every t-stable payoff that is not an optimal t-stable payoff is necessarily dominated by some optimal t-stable payoff.

Proposition 6 $\sum_N u_j = \sum_N w_j > \sum_N z_j$ *for all* $u, w \in S^*$ *and* $z \in S\backslash S^*$.

Proof See the Appendix. □

We note that Proposition 6 provides properties of S^* that are similar to defining properties of the von Neumann-Morgenstern stable set (von Neumann and Morgenstern 2944): two payoff vectors in S^* do not dominate one another, and any payoff vector in $S\backslash S^*$ is dominated by some element of S^*. However, payoffs in S^* are not necessarily efficient. Hence the set of optimal t-stable payoffs is not a von Neumann-Morgenstern stable set.

After providing properties of the optimal t-stable payoffs, we analyze the matchings compatible with optimal t-stable outcomes. Proposition 6 ensures that any such matching induces the maximum surplus out of the total surplus that can be achieved in any t-stable outcome. We will refer to a matching compatible with an optimal t-stable payoff as quasi-optimal. Hence, similar to Proposition 2(b) (which refers to the set of stable outcomes), if (x, u) is an optimal t-stable outcome, then x is a quasi-optimal matching. Proposition 7 states for the set of optimal t-stable outcomes, the property similar to part (a) of Proposition 2(a). It asserts that every quasi-optimal matching is compatible with any optimal t-stable payoff. Therefore, the set of optimal t-stable outcomes is the Cartesian product of the set of optimal t-stable payoffs and the set of quasi-optimal matchings.

Proposition 7 *If* x *is a quasi-optimal matching then it is compatible with any optimal t-stable payoff* u.

Proof See the Appendix. □

To emphasize that the set of optimal t-stable outcomes is the natural extension of the set of stable outcomes when the last set is empty, let us mention that we can use properties of the set of optimal t-stable outcomes, together with Theorem 2, to obtain the corresponding properties of the core as immediate corollaries. In particular, the result that an optimal matching is compatible with any stable payoff (Proposition 2) is a corollary of Proposition 7 and Theorem 2 for the environments where the core is not empty once we realize that the quasi-optimal matchings are optimal if the core exists. And Corollary 1 is just a corollary of that result.

Moreover, given that the relationship between the sets of quasi-optimal matchings and optimal t-stable payoffs is the same as between the set of optimal matchings and the core, we can extend Theorem 2 from payoffs to outcomes. Corollary 2 states this result, whose proof is immediate.

Corollary 2 *The set of core outcomes is non-empty if and only if it coincides with the set of optimal t-stable outcomes.*

Our final result will be instrumental for the dynamic partnership formation process that we will propose in Sect. 5. We show in Corollary 3 that not only there is an optimal t-stable payoff vector whose sum of payoffs is larger than the sum of payoffs of every non-optimal t-stable payoff vector (Proposition 6), but every t-stable outcome that is not optimal can be extended by an optimal t-stable outcome. This result is immediate after Proposition 8:

Proposition 8 *Let (x, u) and (y, w) be t-stable outcomes. Suppose $w > u$. Then (y, w) is an extension of (x, u).*

Proof See the Appendix. □

Therefore, an outcome is optimal t-stable if and only if it cannot be extended by another t-stable outcome, as Corollary 3 states.

Corollary 3 *The set of optimal t-stable outcomes equals the set of t-stable outcomes without an extension.*

Proof Let (x, u) be an optimal t-stable outcome. Then (x, u) cannot have any extension, as stated in Definition 13. The other direction is immediate from Propositions 6 and 8. □

An implication of Corollary 3 is that, in environments where $a_{ij} > 0$ for all $i, j \in N$, $i \neq j$, a matching x is quasi-optimal if and only if $T(x)$ is maximal in $\{T(y) : (y, u)$ is a t-stable outcome$\}$.

5 A dynamic partnership formation process

The intuitive idea of an optimal t-stable outcome is that it corresponds to an outcome that we can expect to occur in an idealized environment where players only engage in cooperation if, once they reach an agreement, they are certain that more favorable options will not be available elsewhere. Indeed, the properties of the t-stable outcomes allow us to envision a dynamic and finite partnership formation process of t-stable outcomes that ends with an optimal t-stable outcome. At every step of this process, some current unmatched agents work among themselves to form partnerships and to split the gains obtained in these partnerships. The process makes sure that, given the current partnerships and payoffs, the matched agents are not interested in trading with any other agent. Because of the properties concerning the extensions of t-stable outcomes (Proposition 6 and Corollary 3) and the fact that the optimal t-stable payoffs extend the t-stable payoffs (Propositions 6 and 8), this process always ends in an optimal t-stable outcome, providing further support for the optimal t-stable outcomes as a natural solution concept for the one-sided assignment game.

To describe the sequential process, consider any t-stable outcome (x, u) such that $T(x) \neq \varnothing$. Define

$$\begin{aligned} A(x,u) \equiv \{&(x^p, u^p) \text{ is t-stable} : T(x^p) \subseteq T(x), \\ &\text{and } x^p(j) = x(j) \text{ and } u_j^p = u_j \text{ for all } j \in T(x^p)\}. \end{aligned}$$

$A(x, u)$ is the set of t-stable outcomes in which the pairs that match do it according to x and have the same payoffs as u.

Denote $B_1(x,u) \equiv \{S^p \subseteq T(x) : S^p \neq \varnothing \text{ and } S^p = T(x^p) \text{ for some } (x^p, u^p) \in A(x,u)\}$. That is, $B_1(x, u)$ is the set of the non-empty stable active coalitions of the t-stable outcomes in $A(x, u)$. The set $B_1(x, u)$ is non-empty since $T(x) \in B_1(x, u)$. Furthermore, $B_1(x, u)$ is finite and is endowed with the partial order defined by the set inclusion relation. Then $B_1(x, u)$ has a minimal element; that is, there exists some coalition that does not have any sub-coalition in $B_1(x, u)$. Set $D_1(x, u)$ any such coalition[12] and let (x^1, u^1) be the corresponding t-stable outcome in $A(x, u)$, that is, (x^1, u^1) is t-stable. The outcome (x^1, u^1) is a "weak" extension of the t-stable outcome (x^0, u^0). We add the term "weak" because (as in Definition 13) we require $T(x^0) \subsetneq T(x^1)$ and $u_j^1 = u_j^0$ for all $j \in T(x^0)$ but, for convenience, we allow here for the possibility that $u_k^1 \geq u_k^0$ for all $k \in T(z)\backslash T(x)$.[13] Moreover, $D_1(x, u) = T(x^1) \subseteq T(x)$. We define $C_1(x, u) \equiv D_1(x, u)$.

If $D_1(x, u) \neq T(x)$, that is, if $(x^1, u^1) \neq (x, u)$, we denote $B_2(x, u) \equiv \{S^p \subseteq T(x) : S^p = T(x^p) \text{ for some } (x^p, u^p) \in A(x, u) \text{ such that } (x^p, u^p) \text{ is a weak extension of } (x^1, u^1)\}$. The set $B_2(x, u)$ is also non-empty because $T(x) \in B_2(x, u)$. By using similar arguments as above, we obtain the existence of a minimal element of $B_2(x, u)$. Set $D_2(x, u)$ any such coalition and let (x^2, u^2) be the corresponding t-stable outcome in $A(x, u)$. By construction, (x^2, u^2) is a t-stable outcome in $A(x, u)$

[12] There can be several minimal coalitions.

[13] Consider the following trivial example. The set of players is $N = \{1, 2\}$, and the surplus of the only possible partnership is $a_{12} = 0$. The outcome (x^1, u^1), with $x^1(1) = 2$ and $u_1^1 = u_2^1 = 0$, is t-stable, and it is a weak extension, but not an extension, of (x^0, u^0). In this market, both (x^0, u^0) and (x^1, u^1) are optimal t-stable outcomes.

and it weakly extends (x^1, u^1). Also, by defining $C_2(x,u) \equiv D_2(x,u)\backslash D_1(x,u)$, we obtain a partition $\{C_1(x,u), C_2(x,u)\}$ of the set $T(u^2)$ of active players in the t-stable outcome (x^2, u^2).

By continuing this procedure, we construct in the first place, a finite set of disjoint subsets of N, $C_1(x,u), C_2(x,u)..., C_k(x,u)$, that safisfy that $T(x) = C_1(x,u) \cup \ldots \cup C_k(x,u)$. In the second place, we build a finite sequence of t-stable outcomes: $(x^0, u^0), (x^1, u^1), \ldots, (x^k, u^k)$, where $(x^k, u^k) = (x, u)$ and (x^{p+1}, u^{p+1}) weakly extends (x^p, u^p) for all $p = 1, \ldots, k-1$. Moreover, the sequence is "maximal," in the sense that there is no t-stable outcome that weakly extends (x^p, u^p) and that is weakly extended by (x^{p+1}, u^{p+1}), for all $p = 0, \ldots, k-1$. We refer to any such sequence as a *t-stable sequence*.

We can describe the sequential process by the t-stable sequence generated by the partition $\{C_1(x,u), \ldots, C_k(x,u)\}$ of $T(x)$ as follows. It starts with the t-stable outcome (x^0, u^0), where all the agents are inactive. In the first step of the process, the players in $C_1(x,u)$ form stable trades that yield (x^1, u^1). Then, in the second step, the players in $C_2(x,u)$ form additional stable trades, which yield (x^2, u^2), and so on. At every step p, no inactive agent is willing to pay any agent i in $\cup_{q=1,\ldots,p} C_q(x,u)$ more than u_i^p. Thus, at any step p, the current outcome (x^p, u^p) is t-stable and is a weak extension of the current outcome (x^{p-1}, u^{p-1}).

We state the results from the previous discussion in Proposition 9.

Proposition 9 *For any t-stable outcome (x,u) with $T(x) \neq \varnothing$ there exists a partition $\{C_1(x,u), \ldots, C_k(x,u)\}$ of $T(x)$ such that a t-stable sequence $(x^p, u^p)_{p=1,\ldots,k}$, with $(x^k, u^k) = (x,u)$, is formed where at each step p only the players in $C_p(x,u)$ interact among themselves to stablish new stable trades.*

If the t-stable outcome (x,u) is not an optimal t-stable outcome, then there is another t-stable outcome that extends (x,u); hence, there is a t-stable sequence whose first steps are $(x^p, u^p)_{p=1,\ldots,k}$ but that continues to reach another t-stable outcome. Hence, the limits of the t-stable sequences are necessarily optimal t-stable outcomes. The process cannot continue once any such optimal t-stable outcome is reached because either no interaction can benefit the agents involved, in which case the core is reached, or any new interaction leads to a set of matched agents which is not internally stable since some of these active agents would have an incentive to switch partners and reach other agreements.

We illustrate one sequential process for the market introduced in Example 2.

Example 2 continued One possible path of the partnership formation process for Example 2 (where $N = \{1, 2, 3, 4, 5, 6, 7\}$, $a_{12} = 3$, $a_{34} = 2$, $a_{13} = 1$, $a_{56} = a_{57} = a_{67} = 1$, and $a_{ij} = 0$ for the other partnerships) is the following.

Step 0/ At the initial step, everyone is unmatched; that is, the oucome is (x^0, u^0).

Step 1/ Players 1 and 2 realize that they may reach an agreement and share the payoff of 3. However, not every sharing of the payment is "reasonable" : if player 1 receives a payoff lower than 1, then they both know that at some point, player 1 can partner with player 3, since $a_{13} = 1$. Therefore, they agree on a sharing that they are certain to keep whatever happens in the future; for instance, player1 can get 2, and player 2 gets 1. In this case, they form the t-stable outcome (x^1, u^1) with $x^1(1) = 2$ and $x^1(i) = i$ for $i = 3, \ldots, 7$; $u_1^1 = 2$, $u_2^1 = 1$, and $u_i^1 = 0$ otherwise.[14]

Step 2/ Players 3 and 4 can follow a similar logic to agree on a sharing of $a_{34} = 2$. They can, for instance, keep 1 each player, knowing that more favorable options will be available elsewhere for neither of them. This way, we reach the t-stable outcome (x^2, u^2) with $x^2(1) = 2, x^2(3) = 4$, and $x^2(i) = i$ for $i = 5, 6, 7$; $u_1^2 = 2, u_2^2 = u_3^2 = u_4^2 = 1$, and $u_i^2 = 0$ otherwise. The process stops at this step (i.e., at this optimal t-stable outcome) because the active players (players 1, 2, 3, and 4) cannot do better, and the outcome cannot be extended to other players in a "stable" way.

The process in the previous example can be generalized for an arbitrary one-sided assignment game. In markets where every partnership creates worth (i.e., $a_{ij} > 0$ for all $i, j \in N$ with $i \neq j$), then every weak extension is an extension, and the set of the limits of the t-stable sequences coincides with the set of the optimal t-stable outcomes. Indeed, on the one hand, we can construct a t-stable sequence for any t-stable outcome, particularly for an optimal one. On the other hand, a t-stable sequence reaches its limit only if it does not have a weak extension (that is, an extension), which, by Corollary 3, means that the limit is an optimal t-stable outcome.

In arbitrary markets, where some partnerships may not create worth, some optimal t-stable outcomes may have a weak extension (see the example in footnote 13). Then the set of outcomes generated by the process coincides with the set of the optimal t-stable outcomes with the largest number of partnerships. They correspond to the optimal t-stable outcomes (x, u) where u is any optimal t-stable payoff and x is a quasi-optimal matching with the largest set of active players.[15] We formally state this result in Proposition 10.

Proposition 10 *The set of the limits of the t-stable sequences coincides with the set of the optimal t-stable outcomes whose quasi-optimal matching has the largest set of active players. Moreover, in markets (N, a) where $a_{ij} > 0$ for all $i, j \in N$ with $i \neq j$, then the set of the limits of the t-stable sequences coincides with the set of the optimal t-stable outcomes.*

We should emphasize that the dynamic partnership formation process described is not an algorithm to find the set of optimal t-stable outcomes. Also, we do not provide a history behind the coalitional interactions among individuals that would lead to the

[14] We note that players 3 and 4 could also have reached an agreement at this step instead of players 1 and 2 (as we stated is footnote 12 , there can be several minimal coalitions $D_q(x, u)$ at step q; hence, there can be several distinct sets $C_q(x, u)$). That is, another possible path of the partnership formation process is that players 3 and 4 reach a partnership at step 1 (and players 1 and 2 reach an agreement at step 2). However, no pair in the subset {5, 6, 7} could have done it because for every possible agreement between, say, players 5 and 6, at least one of them will eventually look for a different partnership with player 7. Any such outcome would not be t-stable.

[15] We say that a quasi-optimal matching x has the largest set of active players if there is no other quasi-optimal matching whose set of active players includes the set of active players of x. We could have used extensions instead of weak extensions in the construction of the sets $B_k(x, u)$ and $C_k(x, u)$. In this case, the outcomes generated by the process would be the optimal t-stable outcomes whose quasi-optimal matchings have the smallest set of active players (a quasi-optimal matching x has the smallest set of active players if there is no other quasi-optimal matching whose set of active players is included in the set of active players of x).

formation of an optimal t-stable outcome. The process only postulates a plausible theory according to which every optimal t-stable outcome may arise. It is based on the property that, for any optimal t-stable outcome, there is a coalitional structure that points to a possible ordering in the formation of these coalitions that ends in that outcome. Moreover, this coalitional structure naturally defines a dynamic procedure where players only trade under the assumption of optimal cooperative behavior since they are sure that any agreement reached will be maintained even if new partnerships are formed later. Only optimal t-stable outcomes occur at the end of such a partnership formation process (this does not imply that every such outcome necessarily occurs). Thus, the process justifies the concept of optimal t-stability as the natural cooperative solution concept for this market.

6 Conditions for a non-empty core

In this section, we further support our approach. We show that the relationship between the core and the set of optimal t-stable payoffs stated in Theorem 2 constitutes a valuable tool to establish conditions under which the core is non-empty. We also provide some interesting properties concerning the blocking pairs of optimal t-stable outcomes.

First, Theorem 3 uses the properties of the set of optimal t-stable payoffs to present a necessary and sufficient condition for the core to be non-empty based on the examination of optimal t-stable payoffs.

Theorem 3 *The set of stable outcomes is non-empty if and only if every optimal t-stable payoff is PO feasible.*

Proof Suppose first that the set of stable outcomes is non-empty and let $u \in S^*$. Theorem 2 implies that $u \in C$, so u is PO feasible.

In the other direction, take an optimal t-stable payoff u, which is also PO feasible, and let x be a t-stable matching compatible with u. Then, $\sum_N u_i = \sum_x (u_i + u_j) = \sum_x a_{ij}$. Since u is PO feasible, then $v(N) = \sum_N u_i$. Therefore, $\sum_x a_{ij} = v(N)$, so x is an optimal matching. Lemma 7 (see Appendix) then implies that $u \in C$, so $C \neq \varnothing$. Hence, the proof is complete. □

Theorem 3 can be linked to the dynamic partnership formation process described in Sect. 5 to state that the core is not empty if and only if, for any outcome (x, u) of the process, $\sum_N u_i = v(N)$.

Our final theorem (Theorem 4) provides a necessary and sufficient condition for the emptiness of the core based on the idea of "non-solvable blocking pairs." Before stating and proving this theorem, we provide two interesting intermediary results.

Proposition 11 provides another feature that is shared by all optimal t-stable outcomes. It also helps us better understand the structure of the optimal t-stable and that of the t-stable outcomes. It states that if an optimal t-stable outcome is not stable, then the set of pairs of blocking agents is the same for every optimal t-stable outcome. Moreover, each of those pairs also blocks any t-stable outcome. This result implies, in particular, that if an agent is unmatched at some optimal t-stable outcome, but he is matched with zero payoff at another optimal t-stable outcome, then that agent will never be part of a blocking pair in an optimal t-stable outcome. And since every t-stable outcome is extended by an optimal t-stable outcome, any blocking agent of an optimal t-stable outcome is unmatched at any t-stable outcome.

Proposition 11 *Let* (x, u) *be an optimal t-stable outcome that is not stable and let* $\{j, k\}$ *be a blocking pair for* (x, u)*. Then,* $\{j, k\}$ *blocks* (y, w)*, for any t-stable outcome* (y, w)*. In particular,* $\{j, k\} \subseteq U(y)$*, for any t-stable outcome* (y, w)*.*

Proof Notice first that, given that $\{j, k\}$ blocks (x, u), it is the case that j and k are unassigned at x, so $0 = u_j + u_k < a_{jk}$. Moreover, j and k have a zero payoff at any optimal t-stable outcome, under Proposition 5 (even if j or k were matched in an optimal t-stable outcome, they would obtain a zero payoff). Therefore, the sum of the payoffs of j and k in any optimal t-stable outcome is less than a_{jk}, so $\{j, k\}$ blocks any optimal t-stable outcome. Now, use Propositions 6 and 8 to get that any t-stable (y, w) which is not optimal t-stable is extended by some optimal t-stable outcome, so $\{j, k\}$ blocks any t-stable (y, w). In particular, j and k are unassigned at y. Hence, the proof is complete. □

Some of the blocking pairs of a t-stable outcome "vanish" along the partnership formation process that we have described at the end of Sect. 5, in the sense that they do not block some t-stable outcomes that extend the original t-stable outcome. Other blocking pairs "persist" along the process as they block all the extensions of the original outcome, including the optimal t-stable outcomes that can be obtained in the last term of the sequences. As we will show in Theorem 4, the last type of blocking pairs play a fundamental role in the emptiness of the core. We will call them "non-solvable blocking pairs."

Definition 14 Let (x, u) be a t-stable outcome and let $\{i, j\} \subseteq U(x)$, with $a_{ij} > 0$ (i.e., $\{i, j\}$ is a blocking pair). We say that $\{i, j\}$ is a *non-solvable blocking pair for* (x, u) if either $u \in S^*$ or $\{i, j\} \subseteq U(x')$ for every extension (x', u') of (x, u). Also, we say that $\{i, j\}$ *is a non-solvable blocking pair* if it is a non-solvable blocking pair for some t-stable outcome (x, u).

Therefore, since $a_{ij} > 0$, if $\{i, j\}$ is a non-solvable blocking pair for (x, u), then $\{i, j\}$ blocks every extension of (x, u), if any. In this case, $\{i, j\}$ also blocks the optimal t-stable outcome which extends (x, u), and that optimal t-stable outcome is corewise-unstable which, under Theorem 2, implies $C = \varnothing$. In fact, every blocking pair $\{i, j\}$ of an optimal t-stable outcome is a non-solvable blocking pair of all t-stable outcomes. This is because, under Proposition 11, the pair $\{i, j\}$ blocks every t-stable outcome (including all optimal t-stable outcomes). Thus, $\{i, j\}$ is a non-solvable blocking pair of every t-stable outcome. Hence, *the set of non-solvable blocking pairs of a given t-stable outcome is the same as that of every t-stable outcome* and, in particular, it coincides with the set of blocking pairs of any optimal t-stable outcome. These conclusions are formalized in the following results.

Proposition 12 *Let* (x, u) *be a t-stable outcome and let* $\{i, j\}$ *be a non-solvable blocking pair for* (x, u)*. Then,* $\{i, j\}$ *is a non-solvable blocking pair of every t-stable outcome.*

Proof As stated in Definition 14, $\{i, j\}$ is a blocking pair of the optimal t-stable outcome that extends (x, u), in case $u \notin S^*$. Then, in any case, $\{i, j\}$ is a blocking pair of an optimal t-stable outcome. According to Proposition 11, $\{i, j\}$ is a blocking pair of every t-stable outcome, and is so for every extension of any t-stable outcome. Definition 14 then implies that $\{i, j\}$ is a non-solvable blocking pair of every t-stable outcome. Hence, the proof is complete. □

We can now state the theorem that provides the conditions for the non-emptiness of the core.

Theorem 4 *The following conditions are equivalent:*

(i) $C = \varnothing$;
(ii) every t-stable outcome has a non-solvable blocking pair;
(iii) there is a t-stable outcome that has a non-solvable blocking pair.

Proof Suppose $C = \varnothing$. Let (y, w) be an optimal t-stable outcome. Since $C = \varnothing$ we have that (y, w) is corewise-unstable according to Theorem 2. Let $\{i, j\} \subseteq U(y)$ with $a_{ij} > 0$. It follows from Proposition 11 that $\{i, j\}$ blocks every t-stable outcome, in particular it blocks every extension of any t-stable outcome, if any. Hence, under Definition 14, $\{i, j\}$ is a non-solvable blocking pair of every t-stable outcome. Then (i) implies (ii). Clearly, (ii) implies (iii).

Now, let (x, u) be a t-stable outcome and suppose $\{i, j\}$ is a non-solvable blocking pair for (x, u). As shown in the proof of Proposition 12, $\{i, j\}$ is a blocking pair of an optimal t-stable outcome. Theorem 2 implies that $C = \varnothing$, so (iii) implies (i). Hence, we have completed the proof. □

As we did with Theorem 3, we can link Theorem 4 to the dynamic partnership formation process: Let (x, u) be any outcome of the process. Then the core is not empty if and only if there is no blocking pair of (x, u).

Finally, we notice that from the previous results, we can conclude that $\{i, j\}$ is a non-solvable blocking pair for some t-stable outcome if and only if the pair $\{i, j\}$ blocks every t-stable outcome. Therefore, if two t-stable outcomes have disjoint sets of blocking pairs, the core is non-empty.

7 Concluding remarks

Our paper studies the one-sided assignment game, which is the generalization of the two-sided assignment game of Shapley and Shubik (1971) to the case where any two agents can form a partnership. It provides a new point of view about stability through the concepts of tradewise-stable outcome and constrained-optimal tradewise-stable outcome.

Tradewise-stable outcomes capture some notion of internal stability: In a tradewise-stable outcome, a matched agent cannot block the situation by deviating with another matched or unmatched agent. In that sense, the set of matched agents (the members of the “club of active agents”) is in a stable situation as none of its members can deviate. The properties of the set of tradewise-stable outcomes allow us to propose a dynamic "partnership formation" process. The set of partners enlarges at each step of the process, but the payoff of the old partners does not change with the arrival of new members. At the end of the process, we always obtain an outcome that is constrained-optimal in the set of t-stable outcomes. The process suggests that when there is no stable outcome, a constrained-optimal t-stable outcome is the most stable outcome that we can expect.

We view the set of constrained-optimal tradewise-stable outcomes as a natural solution concept for the one-sided assignment game. Each of them generates (and they are the only ones that do so) the highest possible total surplus in the set of t-stable outcomes. And, as the previous dynamic process suggests, these outcomes are "as stable as possible," in the sense that any matching involving a larger set of matched agents will necessarily be unstable; the club of active agents would be too large. In fact, the set of constrained-optimal tradewise-stable payoffs coincides with the core when the core is not empty. Thus, the solution concept keeps all the good properties of the core when it exists, but it also provides a prediction for those markets where the core does not exist. Moreover, several of the appealing properties of the core, when it is non-empty, are extended to the set of Pareto-optimal outcomes.

An open question left for future research is an axiomatic characterization of the set of optimal t-stable payoffs in the one-sided assignment game. Sasaki (1995) and Toda (2005) provide characterizations of the core of the two-sided assignment game using the axioms of Pareto optimality, consistency, and pairwise monotonicity together with some other property, such as individual monotonicity. However, no characterization of the core exists for the one-sided assignment game.[16] Given that the set of t-stable payoffs coincide with the core when the core is non-empty, the characterization of both sets might be obtained through similar axioms.

Bondareva (1963) and Shapley (1967) proved that the core of a transferable utility game is non-empty if and only if the game is balanced. Thus, for the game considered here, the condition that every optimal tradewise-stable payoff is Pareto-optimal feasible is equivalent to balancedness. This suggests the question of whether this equivalence persists in all transferable-utility (TU) games. The answer to this question is not easy. Our results strongly rely on the existence of a feasible matching underlying every feasible outcome. However, players do not necessarily form partnerships in the general TU game. On the other hand, the intuition behind a tradewise-stable outcome is not related to a matching and seems to be quite general: if all "interactions" are made under the premise of optimal behavior, a tradewise-stable outcome results. In subsequent work, Pérez-Castrillo and Sotomayor (2019) consider the extension of the present investigation to the coalitional games with transfers. This extension is not straightforward because our concepts make use of the fact that every feasible outcome is compatible with a feasible matching. The solution of this problem is possible by the identification, for every feasible outcome, of some convenient coalitional structure that, restricted to the one-sided matching model, coincides with a feasible matching.

[16] Klaus and Nichifor (2010) provide some results on the one-sided assignment game. For instance, they show that, for the subset of games where the core is non-empty if a subsolution of the core satisfies consistency and Pareto indifference, it coincides with the core.

8 Appendix

Lemma 1 *The set of t-stable payoffs S is a compact set of $\mathbb{R}^n$.*

Proof The set S is bounded because $0 \leq u_j \leq v(N)$ for all $j \in N$ and for all t-stable payoffs u. To prove that it is also closed, take any sequence $(u^t)_{t=1,2,\ldots}$ of t-stable payoffs, with $u^t \to u$ when t tends to infinity. Since the set of matchings is finite, there is some matching x which is compatible with infinitely many terms of the sequence $(u^t)_{t=1,2,\ldots}$. Denote $(v^t)_{t=1,2,\ldots}$ this subsequence. Then, if $x(j) = k$, $u_j + u_k = lim_{t\to\infty}(v_j^t + v_k^t) = lim_{t\to\infty} a_{jk} = a_{jk}$. Similarly, if $x(j) = j$ then $u_j = lim_{t\to\infty} v_j^t = 0$. Thus, x is compatible with u, so (x, u) is feasible. We claim that if j is matched at x then j is not part of a blocking pair of (x, u). In fact, $u_j + u_k = lim_{t\to\infty}(v_j^t + v_k^t) \geq lim_{t\to\infty} a_{jk} = a_{jk}$ for any $k \in N\backslash\{j\}$, where the inequality holds because (v^t, x) is a t-stable outcome for all t. Therefore, (x, u) is a t-stable outcome, so u is a t-stable payoff. Hence, the set of t-stable payoffs is bounded and closed, so it is compact. □

Lemma 2 *Let A be a non-empty and compact set of $\mathbb{R}^n$, ordered with the partial order relation $\geq$ induced by $\mathbb{R}^n$. Then, the set of maximal elements of A with respect to $\geq$ is a non-empty and compact set of $\mathbb{R}^n$.*

Proof Denote $A^* \equiv \{u \in A : u$ is a maximal element of $A\}$. It is known that every non-empty, compact, and partially ordered set has a maximal element, so $A^* \neq \varnothing$. The set A^* is bounded since A is bounded. To see that A^* is closed, take any sequence of vectors $(u^t)_{t=1,2,\ldots}$, with $u^t \in A^*$ for all t, which converges to some vector u. Suppose, by way of contradiction, that $u \notin A^*$. Then, there exists some vector $w \in A$ such that $w > u$. If this is the case, there is some neighborhood V of the vector u and some integer k such that $u^t \in V$ for all $t \geq k$ and $w > u'$ for all $u' \in V$. In particular, $w > u^k$, which contradicts the assumption that $u^k \in A^*$. Hence, A^* is a compact set of $\mathbb{R}^n$. □

Lemma 3 *Let (x, u) be a t-stable outcome. Then, $v(T(x)) + v(U(x)) = v(N)$.*

Proof Let y be an optimal matching, that is, $v(N) = \sum_y a_{ij}$. Set

$$\alpha \equiv \{\{i, j\} \in y : \{i, j\} \cap T(x) \neq \varnothing \text{ and } \{i, j\} \cap U(x) \neq \varnothing\},$$
$$\beta \equiv \{\{i, j\} \in y : \{i, j\} \subseteq T(x)\}, \text{ and}$$
$$\gamma \equiv \{\{i, j\} \in y : \{i, j\} \cap T(x) = \varnothing\}.$$

Also, denote

$$\alpha_x \equiv \{i \in T(x) : \{i, j\} \in \alpha \text{ for some } j\}$$

and

$$R \text{ is a set of pairs } \{i, j\} \subseteq U(x) \text{ such that } v(U(x)) \equiv \sum_R a_{ij},$$

$$R' \equiv U(x)\backslash \cup_\gamma \{i, j\}, \text{ and}$$

$$R'' \text{ is a set of pairs } \{i, j\} \subseteq T(x) \text{ such that } v(T(x)) \equiv \sum_{R''} a_{ij}.$$

Then, $\sum_{T(x)} u_i = \sum_{\alpha_x} u_i + \sum_{\beta} a_{ij} \geq \sum_{\alpha} a_{ij} + \sum_{\beta} a_{ij}$, where the inequality holds because (x, u) is t-stable, $i \in T(x)$, and $u_{x(i)} = 0$ for all $i \in \alpha_x$, and so $u_i = u_i + u_{x(i)} \geq a_{ix(i)}$ for all $i \in \alpha_x$.

Also, $R' \cup \gamma$ is a partition of $U(x)$, so $v(U(x)) = \sum_R a_{ij} \geq \sum_{\gamma} a_{ij} + \sum_{R'} a_{ii} = \sum_{\gamma} a_{ij}$. Then,

$$\begin{aligned} v(N) = \sum_{y} a_{ij} &= \left(\sum_{\alpha} a_{ij} + \sum_{\beta} a_{ij}\right) + \sum_{\gamma} a_{ij} \leq \sum_{T(x)} u_i + v(U(x)) \\ &= v(T(x)) + v(U(x)), \end{aligned}$$

where the last equality uses that $T(x)$ is a stable coalition. Therefore,

$$v(N) \leq v(T(x)) + v(U(x)). \tag{3}$$

On the other hand, y is an optimal matching, so $v(N) = \sum_y a_{ij} \geq \sum_{R \cup R''} a_{ij} = v(T(x)) + v(U(x))$. Then,

$$v(N) \geq v(T(x)) + v(U(x)). \tag{4}$$

Hence, $v(T(x)) + v(U(x)) = v(N)$ and the proof is complete. □

Proof of Proposition 4 According to Proposition 3, the set of active partnerships of x is part of some optimal matching. Therefore, there is some optimal matching z such that $z(i) = x(i)$ for all $i \in T(x)$. Let (z, w) be any stable outcome. Construct the outcome (z, u^*) such that $u_i^* = u_i$ for all $i \in T(x)$ and $u_i^* = w_i$ for all $i \in N \backslash T(x)$. The outcome (z, u^*) is feasible. We claim that $u^* \in C$. In fact, suppose $\{i, j\}$ blocks u^*. Then, $u_i^* + u_j^* < a_{ij}$. Notice that, by construction, $u^* \geq u$, so $u_i + u_j < a_{ij}$. Since x is t-stable, we must have that $\{i, j\} \subseteq N \backslash T(x)$. On the other hand, the stability of w implies that $u_i^* + u_j^* = w_i + w_j \geq a_{ij}$, which contradicts the assumption that $\{i, j\}$ blocks u^*. Then, u^* does not have any blocking pair.

The individual rationality of (z, u^*) is immediate after the individual rationality of (x, u) and (z, w). According to Definition 6, it remains to show that $v(N) = \sum_N u_i^*$. Write:

$$\begin{aligned} v(N) &\leq \sum_{N} u_i^* = \sum_{T(x)} u_i + \sum_{N \backslash T(x)} w_i = v(T(x)) + \sum_{N \backslash T(x)} w_i \leq \sum_{T(x)} w_i + \sum_{N \backslash T(x)} w_i \\ &= \sum_{N} w_i = v(N), \end{aligned} \tag{5}$$

where in the first inequality, we use that u^* does not have any blocking pair, in the second equality, we use Remark 3, and the second inequality follows from the stability of w. Then, the inequalities in (5) must be equalities, so $v(N) = \sum_N u_i^*$. Therefore, we have proved that (z, u^*) is a stable outcome.

To see that (z, u^*) extends (x, u), use that $u_i^* \geq u_i$ for all $i \in N$. Given that (z, u^*) is a stable outcome and that (x, u) is unstable, we have that $\{j \in N : u_j^* > u_j\} \neq \varnothing$. On the other hand, since $u_j^* = u_j$ for all $j \in T(x)$, it follows that $\{j \in N : u_j^* > u_j\} \subseteq N \backslash T(x)$. Then, according to Definition 13, (z, u^*) extends (x, u), and we have proved the proposition. □

The next result is a *Decomposition Lemma* for the set of t-stable outcomes that has similarities with other decomposition lemmas in matching models (see, for instance, Gale and Sotomayor 1985). It states that for any two t-stable outcomes, a player who is matched at both outcomes and obtains a higher payoff in the first is necessarily matched, at both outcomes, to a player who obtains a higher payoff in the second.

Lemma 4 *Let (x, u) and (y, w) be t-stable outcomes. Let $M_u \equiv \{j \in T(y) : u_j > w_j\}$ and $M_w \equiv \{j \in T(x) : w_j > u_j\}$. Then $x(M_u) = y(M_u) = M_w$ and $x(M_w) = y(M_w) = M_u$.*[17]

Proof We first prove that $x(M_u) \subseteq M_w$. Take $j \in M_u$; then j is matched under x since $u_j > w_j \geq 0$. We show by contradiction that $k \equiv x(j)$ is in M_w. Suppose $k \notin M_w$, then

$$a_{jk} = u_j + u_k > w_j + w_k$$

which implies that (j, k) blocks (y, w). However, $j \in M_u$ so it is matched at y, which contradicts that (y, w) is t-stable.

A similar argument leads to $y(M_w) \subseteq Mu$.

Moreover, $x(M_u) \subseteq M_w$ implies $M_u \subseteq x(M_w)$ and $y(M_w) \subseteq M_u$ implies $M_w \subseteq y(M_u)$. Since all the players in M_u and in M_w are matched at x and y, it follows that $|M_u| = |x(M_u)|$, $|M_w| = |y(M_w)|$, $|y(M_u)| = |M_u|$ and $|y(M_u)| = |M_u|$. Therefore,

$$|M_u| = |x(M_u)| \leq |M_w| = |y(M_w)| \leq |M_u|$$

and

$$|M_w| \leq |y(M_u)| = |M_u| \leq |x(M_w)| = |M_w|,$$

which imply $x(M_u) = M_w$, $y(M_w) = M_u$, $y(M_u) = M_w$, and $x(M_w) = M_u$. □

We notice that, in the proof of Lemma 4, it is shown that $u_j > 0$ for all $j \in M_u$ and $w_j > 0$ for all $j \in M_w$. Therefore, we can write $M_u = \{j \in T(x) \cap T(y) : u_j > w_j\}$ and $M_w \equiv \{j \in T(x) \cap T(y) : w_j > u_j\}$.

Proof of Proposition 5 Denote $k^* \equiv x(j^*)$. If $a_{j^*k^*} = 0$, then $u_{j^*} = w_{j^*} = 0$ and $u_{k^*} = 0$. If $k^* \in T(x) \backslash T(y)$ then $w_{k^*} = 0$. Otherwise, we cannot have that $u_{k^*} \neq w_{k^*}$,

[17] The decomposition lemma applies, in particular, to stable outcomes. Then, an immediate consequence of the lemma is a polarization of interests between the partners along the core: If (x, u) and (y, w) are stable outcomes, j is matched to k under x or under y, and $u_j > w_j$, then $w_k > u_k$. This is because both payoffs are compatible with the same optimal matching; therefore, if j is matched to k under (x, u) then j is also matched to k under (x, w), so Lemma 4 applies.

because Lemma 4 would imply that $j^* \in T(x) \cap T(y)$, which would be a contradiction. Therefore, it is always the case that $u_{x(j^*)} = w_{x(j^*)}$.

Suppose now that $a_{j^*k^*} > 0$. Under Lemma 6, $k^* \in T(x) \cap T(y)$. Then, since $j^* \notin T(y)$ we have that $w_{j^*} = 0$. In addition, we cannot have that $u_{k^*} \neq w_{k^*}$, according to Lemma 4 and the assumption that $j^* \in T(x) \backslash T(y)$, so $u_{k^*} = w_{k^*}$. Now, suppose by way of contradiction that $u_{j^*} > 0$. Then, $u_{k^*} < a_{j^*k^*}$. Therefore, $w_{j^*} + w_{k^*} = u_{k^*} < a_{j^*k^*}$, so $\{j^*, k^*\}$ blocks (y, w), which is a contradiction because $k^* \in T(y)$ and (y, w) is a t-stable outcome. Hence, $u_{j^*} = w_{j^*} = 0$ and $u_{x(j^*)} = w_{x(j^*)}$, and the proof is complete. □

Lemma 5 *Let (x, u) be a t-stable outcome which is not optimal t-stable, that is, $u \in S \backslash S^*$. Then there is some optimal t-stable payoff w_u such that $w_u > u$.*

Proof of Lemma 5 Suppose, by way of contradiction, that there is no payoff w in S^* such that $w > u$. Since $u \notin S^*$, there is some $w^1 \in S$ such that $w^1 > u$. Then, by contradiction, $w^1 \notin S^*$, so there is some $w^2 \in S$ such that $w^2 > w^1 > u$. Again, w^2 cannot be in S^*. By repeating this procedure, we obtain an infinite sequence $(w^t)_{t=1,2,\ldots}$ of t-stable payoffs with different terms. On the other hand, there is a finite number of t-stable matchings, so there is some t-stable matching x compatible with infinitely many terms of the sequence. Denote $(v^t)_{t=1,2,\ldots}$ that subsequence. All the members of the subsequence are in S so $\sum_{T(x)} v^t_j = v(T(x))$ for all $t = 1, 2, \ldots$. However, $v^1 < v^2$, so $v(T(x)) = \sum_{T(x)} v^1_j = \sum_N v^1_j < \sum_N v^2_j = \sum_{T(x)} v^2_j = v(T(x))$, which is an absurd. Hence, there is some $w_u \in S^*$ such that $w_u > u$. □

Proof of Proposition 6 First, since S^* is a non-empty and compact set of $\mathbb{R}^n$ (see Theorem 1) and every continuous function defined in a compact set has a maximum in this set, there is some $w^* \in S^*$ such that $\sum_N w^*_j \geq \sum_N w_j$ for all $w \in S^*$. Now use Lemma 5 to get that $\sum_N w^*_j > \sum_N u_j$ for all $u \in S \backslash S^*$. Hence, there is some some optimal t-stable payoff that dominates every non-optimal t-stable payoff.

Second, we show that $\sum_N u_j = \sum_N w_j$ for all $u, w \in S^*$. Let (x, u) and (y, w) be optimal t-stable outcomes. Set

$$
\begin{aligned}
B_1(x) &= \{j \in T(x) \cap T(y) : x(j) \in T(x) \cap T(y)\};\\
B_1(y) &= \{j \in T(x) \cap T(y) : y(j) \in T(x) \cap T(y)\};\\
B_2(x) &= \{j \in T(x) \cap T(y) : x(j) \in T(x) \backslash T(y)\};\\
B_2(y) &= \{j \in T(x) \cap T(y) : y(j) \in T(y) \backslash T(x).
\end{aligned}
$$

Clearly,

$$
T(x) \cap T(y) = B_1(x) \cup B_2(x) = B_1(y) \cup B_2(y). \tag{6}
$$

Under Remark 3,

$$
\sum_{B_1(x)} u_j = v(B_1(x)) \text{ and } \sum_{B_1(y)} w_j = v(B_1(y)). \tag{7}
$$

On the other hand, according to Proposition 5, $u_j = w_j$ for all $j \in B_2(x)$ and $w_j = u_j$ for all $j \in B_2(y)$, so

$$\sum_{B_2(x)} u_j = \sum_{B_2(x)} w_j \text{ and } \sum_{B_2(y)} w_j = \sum_{B_2(y)} u_j. \tag{8}$$

Moreover, Proposition 5 implies that

$$\sum_{T(x)\backslash T(y)} u_j = 0 \text{ and } \sum_{T(y)\backslash T(x)} w_j = 0. \tag{9}$$

Therefore, we can write,

$$\begin{aligned}\sum_N u_j &= \sum_{T(x)} u_j + \sum_{N\backslash T(x)} u_j = \sum_{T(x)} u_j = \sum_{T(x)\cap T(y)} u_j + \sum_{T(x)\backslash T(y)} u_j = \sum_{T(x)\cap T(y)} u_j \\ &= \sum_{B_1(x)} u_j + \sum_{B_2(x)} u_j = v(B_1(x)) + \sum_{B_2(x)} w_j \leq \sum_{B_1(x)} w_j + \sum_{B_2(x)} w_j \\ &= \sum_{T(x)\cap T(y)} w_j = \sum_{T(x)\cap T(y)} w_j + \sum_{T(y)\backslash T(x)} w_j = \sum_{T(y)} w_j \\ &= \sum_{T(y)} w_j + \sum_{N\backslash T(y)} w_j = \sum_N w_j,\end{aligned}$$

where the fourth equality uses (9); the fifth equality follows from (6); the sixth equality follows from (7) and (8); the inequality follows from the fact that $B_1(x) \subseteq T(y)$ and y is a t-stable matching, and so $B_1(x)$ cannot block y; the seventh equality follows from (6); and the eighth equality follows from (9).

Then,

$$\sum_N u_j \leq \sum_N w_j. \tag{10}$$

By reversing the roles between (x, u) and (y, w) in the expression (10) we obtain

$$\sum_N w_j \leq \sum_N u_j. \tag{11}$$

Equations (10) and (11) imply that $\sum_N u_j = \sum_N w_j$ and we have completed the proof. □

Proof of Proposition 7 Let (y, w) be any optimal t-stable outcome. We want to show that $w_i + w_j = a_{ij}$ if $x(i) = j$ and $w_i = 0$ if $i \in U(x)$. Notice that

$$\begin{aligned}\sum_{T(x)} u_i &= \sum_{T(x)} u_i + \sum_{U(x)} u_i = \sum_N u_i = \sum_N w_i = \sum_{T(x)} w_i + \sum_{U(x)} w_i \\ &= \sum_{T(x)} w_i + \sum_{T(y)\backslash T(x)} w_i + \sum_{U(y)\backslash T(x)} w_i = \sum_{T(x)} w_i,\end{aligned}$$

where the third equality is due to Proposition 6; and Proposition 5 was used in the last equality to conclude that $\sum_{T(y)\setminus T(x)} w_i = 0$. Then,

$$\sum_{T(x)} u_i = \sum_{T(x)} w_i. \tag{12}$$

We can write $T(x) = (T(x) \cap T(y)) \cup (T(x)\setminus T(y))$. From Proposition 5 it follows that $\sum_{T(x)\setminus T(y)} u_i = \sum_{T(x)\setminus T(y)} w_i = 0$. Then, $\sum_{T(x)} u_i = \sum_{T(x)\cap T(y)} u_i$ and $\sum_{T(x)} w_i = \sum_{T(x)\cap T(y)} w_i$. Therefore, using (12) we obtain

$$\sum_{T(x)\cap T(y)} u_i = \sum_{T(x)\cap T(y)} w_i. \tag{13}$$

To prove that $w_i + w_j = a_{ij}$ if $x(i) = j$, set $G \equiv \{\{i, j\} \subseteq T(x)\cap T(y) : x(i) = j\}$ and $H \equiv \{\{i, j\} \subseteq T(x) : i \in T(x) \cap T(y), j \in T(x)\setminus T(y)$ and $x(i) = j\}$. Proposition 5 implies that

$$\sum_{H} u_i = \sum_{H} w_i. \tag{14}$$

Moreover, since w is t-stable, we must have that $w_i + w_j \geq a_{ij}$ for all $\{i, j\} \subseteq T(x) \cap T(y)$. Then, in particular,

$$\sum_{G} a_{ij} \leq \sum_{G} (w_i + w_j). \tag{15}$$

Therefore,

$$\begin{aligned}\sum_{T(x)\cap T(y)} u_i &= \sum_{G}(u_i + u_j) + \sum_{H} u_i = \sum_{G} a_{ij} + \sum_{H} w_i \leq \sum_{G}(w_i + w_j) + \sum_{H} w_i \\ &= \sum_{T(x)\cap T(y)} w_i,\end{aligned}$$

where we used (14) in the second equality and (15) in the inequality. According to (13), the inequality must be an equality. Hence, we have proved that $w_i + w_j = a_{ij}$ for all $\{i, j\}$ such that $\{i, j\} \subseteq (T(x) \cap T(y))$ and $x(i) = j$. Next, consider $\{i, j\}$ with $x(i) = j$ such that either $i \in (T(x)\setminus T(y))$ or $j \in (T(x)\setminus T(y))$. Without loss of generality suppose that $i \notin T(y)$. Then, $w_i = 0$ and, under Proposition 5, we have that $u_i = 0$ and $w_j = u_j$, from which follows that $w_i + w_j = u_i + u_j = a_{ij}$, so $w_i + w_j = a_{ij}$.

It remains to show that $w_i = 0$ for all $i \in U(x)$. This is immediate from the fact that if $i \in T(y)\setminus T(x)$, then Proposition 5 implies that $w_i = 0$. Hence, the matching x is compatible with w, and we have completed the proof. □

Proof of Proposition 8 To show that (y, w) is an extension of (x, u), we need to prove that $j \notin T(x)$ for all j such that $w_j > u_j$ (see Definition 13). Consider a player j such that $w_j > u_j$ and suppose, by contradiction, that $j \in T(x)$. Given that $w_j > 0$, we have that $j \in T(y)$. Then, $j \in T(x) \cap T(y)$ and so $j \in M_w \equiv \{j \in T(x) : w_j > u_j\}$. Denote $k \equiv y(j)$. Lemma 4 implies that $k \in M_u \equiv \{j \in T(y) : u_j > w_j\}$. Therefore, $u_k > w_k$, which contradicts the assumption that $w > u$. Hence, (y, w) extends (x, u). $\square$

Lemma 6 *Let (x, u) and (y, w) be optimal t-stable outcomes. Let $j^* \in T(x) \backslash T(y)$ with $a_{j^* x(j^*)} > 0$. Then $x(j^*) \in T(y)$.*

Proof Suppose, by way of contradiction, that $x(j^*) \in T(x) \backslash T(y)$. Denote $A \equiv \{t \in T(x) \backslash T(y) : x(t) \in T(x) \backslash T(y)\}$. We have that $j^* \in A$, so $A \neq \varnothing$.

We first show that

$$w_{q^*} + u_{t^*} = u_{t^*} < a_{q*t^*} \text{ for some } t^* \in A \text{ and } q^* \in N \backslash (T(y) \cup A). \tag{16}$$

By way of contradiction, suppose that there are no such players t^* and q^*. Then, either $N \backslash (T(y) \cup A) = \varnothing$ or (given that $w_q = 0$ for all $q \notin T(y)$) $u_t \geq a_{qt}$ for all $q \in N \backslash (T(y) \cup A)$ and $t \in A$. In any case, we can construct the outcome (w', y') as follows: the matching y' agrees with x on A, and it agrees with y on $N \backslash A$; hence, $T(w') = T(y) \cup A$; the payoff vector satisfies $w'_j = w_j$ for all $j \in T(y)$, $w'_j = u_j$ for all $j \in A$ and $w'_j = 0$ for all $j \in N \backslash (T(y) \cup A)$. Since (x, u) is t-stable, there is no pair blocking (w', y') among the agents of A. Because (y, w) is t-stable, there is no blocking pair formed by two agents in $T(y)$ or an agent in A and an agent in $T(y)$. Finally, if $N \backslash (T(y) \cup A) \neq \varnothing$ then first, no blocking pair exists between an agent in $T(y)$ and an agent in $N \backslash (T(y) \cup A)$ because w' coincides with w for these agents and (y, w) is t-stable and second, by using the contradiction assumption, $0 + w'_t = u_t \geq a_{qt}$ for all $t \in A$ and all $q \in N \backslash (T(y) \cup A)$. Therefore, the outcome (w', y') is t-stable. Since $a_{j^* x(j^*)} > 0$, it follows that either $u_{j^*} > 0$ or $u_{x(j^*)} > 0$. Hence, the outcome (w', y') is an extension of (y, w), which is a contradiction because (y, w) is an optimal t-stable outcome.

Once we have shown that there exist some $t^* \in A$ and $q^* \in N \backslash (T(y) \cup A)$ such that $w_{q^*} + u_{t^*} = u_{t^*} < a_{q^* t^*}$, we claim that such q^* necessarily satisfies $q^* \in T(x)$ and $u_{q^*} > 0$. Otherwise, $u_{q^*} = 0$, in which case $u_{q^*} + u_{t^*} < a_{q^* t^*}$ by (16), and then $\{q^*, t^*\}$ would block (x, u), which is not possible because $t^* \in T(x)$ and (x, u) is a t-stable outcome. Therefore, $q^* \in T(x) \backslash (T(y) \cup A)$. Since $q^* \notin A$, we must have that $p \equiv x(q^*) \in T(x) \cap T(y)$, so

$$u_p = w_p \tag{17}$$

because otherwise we should have that $q^* \in T(x) \cap T(y)$ according to Lemma 4, which would be a contradiction. Furthermore, the fact that $u_{q^*} > 0$ implies that $u_p < a_{pq^*}$, and so $w_p < a_{pq^*}$ by (17), which implies that $\{q^*, p\}$ blocks (y, w) (because $w_{q^*} = 0$), which is a contradiction since $p \in T(y)$. Hence, $x(j^*) \in T(x) \cap T(y)$, and the proof is complete. $\square$

Lemma 7 *Let (x, u) be a t-stable outcome. Suppose x is optimal. Then u is a stable outcome.*

Proof Denote R a set of pairs $\{i, j\} \subseteq T(x)$ such that $v(T(x)) \equiv \sum_R a_{ij}$. Since x is optimal, then

$$v(N) = \sum_x a_{ij} = \sum_{x|_{T(x)}} a_{ij} + \sum_{x|_{U(x)}} a_{ij} = \sum_{x|_{T(x)}} a_{ij} = v(T(x)),$$

where the last equality follows from Proposition 3. Then, under Lemma 3, we have that $v(U(x)) = 0 = \sum_{U(x)} u_i$ so, as stated in Remark 4, there is no blocking pair in $U(x)$, which implies that (x, u) is a stable outcome. Hence, we have completed the proof. □

Funding Open Access Funding provided by Universitat Autonoma de Barcelona.

References

Andersson, T., Gudmundsson, J., Talman, D., Yang, Z.: A competitive partnership formation process. Games Econ. Behav. **86**, 165–177 (2014)

Bondareva, O.N.: Some applications of linear programming methods to the theory of cooperative games. Problemy Kybernetiki **10**, 119–139 (1963). (**[in Russian]**)

Chiappori, P.-A., Galichon, A., Salanié, B.: The roommate problem is more stable than you think. CESifo WP 4676 (2014)

d'Aspremont, C., Jacquemin, A., Gabszewicz, J.J., Weymark, J.A.: On the stability of collusive price leadership. Can. J. Econ. **16**, 17–25 (1983)

Demange, G., Gale, D.: The strategy structure of two-sided matching markets. Econometrica **53**, 873–888 (1985)

Demuynck, T., Herings, P.J.J., Saulle, R.D., Seel, C.: The myopic stable set for social environments. Econometrica **87**, 111–138 (2019)

Eriksson, K., Karlander, J.: Corewise-stable matching in a common generalization of the marriage and assignment models. Discrete Math. **217**, 135–156 (2000)

Gale, D., Shapley, L.: College admissions and the stability of marriage. Am. Math. Monthly **69**, 9–15 (1962)

Gale, D., Sotomayor, M.: Some remarks on the corewise-stable matching problem. Discrete Appl. Math. **11**, 223–32 (1985)

Herings, P.J.J., Kóczy, L.Á.: The equivalence of the minimal dominant set and the myopic stable set for coalition function form games. Games Econ. Behav. **127**, 67–79 (2021)

Herings, P.J.J., Mauleon, A., Vannetelbosch, V.: Farsightedly stable networks. Games Econ. Behav. **67**, 526–541 (2009)

Herings, P.J.J., Mauleon, A., Vannetelbosch, V.: Stability of networks under horizon-K farsightedness. Econ. Theory **68**, 177–201 (2019)

Klaus, B., Nichifor, A.: Consistency in one-sided assignment problems. Soc. Choice Welf **35**, 415–433 (2010)

Klaus, B., Payot, F.: Paths to stability in the assignment problem. J. Dyn. Games **2**, 257–287 (2015)

Kóczy, L.Á., Lauwers, L.: The minimal dominant set is a non-empty core-extension. Games Econ. Behav. **61**, 277–298 (2007)

Luo, C., Mauleon, A., Vannetelbosch, V.: Network formation with myopic and farsighted players. Econ. Theory **71**, 1283–1317 (2021)
Pérez-Castrillo, D., Sotomayor, M.: Constrained-optimal tradewise-stable outcomes in coalitional games with transferable payoff: a solution concept weaker than the core, mimeo (2019)
Roth, A.: On the allocation of residents to rural hospitals: a general property of two-sided matching markets. Econometrica **54**, 425–27 (1986)
Roth, A., Sotomayor, M.: Two-sided Matching: A Study in Game-theoretic Modeling and Analysis. Econometric Society Monographs, 2nd ed., vol. 18, Cambridge University Press (1990)
Roth, A., Vande Vate, J.H.: Random paths to stability in two-sided matching. Econometrica **58**, 1475–1480 (1990)
Sasaki, H.: Consistency and monotonicity in assignment problems. Int. J. Game Theory **24**, 373–397 (1995)
Shapley, L.S.: On balanced sets and cores. Nav. Res. Logist. Q **14**, 453–460 (1967)
Shapley, L., Scarf, H.: On cores and indivisibility. J. Math. Econ. **1**, 23–8 (1974)
Shapley, L., Shubik, M.: The assignment game I: the core. Int. J. Game Theory **1**, 111–130 (1971)
Sotomayor, M.: A non-constructive elementary proof of the existence of corewise-stable marriages. Games Econ. Behav. **13**, 135–137 (1996)
Sotomayor, M.: Three remarks on the many-to-many stable matching problem. Math. Soc. Sci. **38**, 55–70 (1999)
Sotomayor, M.: Existence of stable outcomes and the lattice property for a unified matching market. Math. Social Sci. **39**, 119–132 (2000)
Sotomayor, M.: Implementation in the many-to-many matching market. Games Econ. Behav. **46**, 199–212 (2004)
Sotomayor, M.: On the Core of the One-sided Assignment Game. WP Universidade de São Paulo (2005a)
Sotomayor, M.: An elementary non-constructive proof of the non-emptiness of the core of the housing market of Shapley and Scarf. Math. Soc. Sci. **50**, 298–303 (2005b)
Sotomayor, M.: The Role Played by Simple Outcomes in Coalition Formation Process of Core Outcomes. WP Universidade de São Paulo (2009a)
Sotomayor, M.: Coalitional interactions leading to new solution concepts in the time-sharing assignment game, mimeo (2009b)
Sotomayor, M.: The Pareto-stability concept is a natural solution concept for discrete matching markets with indifferences. Int. J. Game Theory **40**, 631–644 (2011)
Sotomayor, M.: Simple Allocations in the Assignment Game. WP Universidade de São Paulo (2018)
Sotomayor, M.: Matching with Stable Trades, mimeo (2019)
Talman, D., Yang, Z.: A model of partnership formation. J. Math. Econ. **47**, 206–212 (2011)
Toda, M.: Consistency and monotonicity in assignment problems. Games Econ. Behav. **53**, 248–261 (2005)
von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1944)
Wu, Q., Roth, A.: The lattice of envy-free matchings. Games Econ. Behav. **109**, 201–211 (2018)