Strategy-Proofness of Worker-Optimal Matching with Continuously Transferable Utility*

Ravi Jagadeesan
Department of Mathematics
Harvard University

Scott Duke Kominers
Harvard Business School and Department of Economics
Harvard University

Ross Rheingans-Yoo
Center of Mathematical Sciences and Applications
Harvard University

August 25, 2017

Abstract

We give the first direct proof of one-sided strategy-proofness for worker–firm matching under continuously transferable utility. A new "Lone Wolf" theorem (Jagadeesan et al. (2017)) for settings with transferable utility allows us to adapt the method of proving one-sided strategy-proofness that is typically used in settings with discrete transfers.

JEL Classification: C78, D44, D47, D82

Keywords: Matching, Strategy-proofness, Lone Wolf Theorem, Rural Hospitals Theorem, Mechanism design

*This research was conducted while Jagadeesan and Rheingans-Yoo were Economic Design Fellows at the Harvard Center of Mathematical Sciences and Applications (CMSA). The authors thank John William Hatfield, Shengwu Li, Alvin E. Roth, Jan Christoph Schlegel, Alexander Teytelboym, an anonymous referee, and the editor, David K. Levine, for helpful comments. Jagadeesan gratefully acknowledges the support of a Harvard Mathematics Department travel grant. Kominers gratefully acknowledges the support of National Science Foundation grant SES-1459912 and the Ng Fund of the CMSA.

1 Introduction

One key reason the Gale–Shapley (1962) deferred acceptance mechanism has been attractive for practical applications is *one-sided strategy-proofness*—the mechanism is dominant-strategy incentive compatible for one side of the market (Dubins and Freedman (1981); Roth (1982); Martínez et al. (2004)).[1,2] The one-sided strategy-proofness result for deferred acceptance has been extended to nearly all settings with discrete transfers (or other discrete contracts; see Roth and Sotomayor (1990); Hatfield and Milgrom (2005); Hatfield and Kojima (2009, 2010); Hatfield and Kominers (2012, 2015); Hatfield, Kominers, and Westkamp (2016)). In contrast, for most settings with continuous transfers, strategy-proofness results for deferred acceptance were not known until recently.[3] Indeed, it appears that until the work of Hatfield, Kojima, and Kominers (2017), one-sided strategy-proofness of deferred acceptance in the presence of continuously transferable utility was known only for one-to-one matching markets (Demange (1982); Leonard (1983); Demange and Gale (1985); Demange (1987)).[4]

The now-standard proof of one-sided strategy-proofness by Hatfield and Milgrom (2005) relies on a classic matching-theoretic result called the "Lone Wolf Theorem" (McVitie and Wilson (1970); Roth (1984a, 1986)). Unfortunately, even non-generic indifferences can undermine the Lone Wolf Theorem—and discrete choice in the presence of continuous transfers necessarily involves indifferences. Consequently, one-sided strategy-proofness results have heretofore been difficult to derive in matching settings with continuously transferable utility.[5]

[1]Strategy-proofness has been key to the adoption of deferred acceptance in both medical resident matching (see, e.g., Roth (1984a); Roth and Peranson (1999)) and school choice (see, e.g., Balinski and Sönmez (1999); Abdulkadiroğlu and Sönmez (2003); Abdulkadiroğlu, Pathak, and Roth (2005); Abdulkadiroğlu, Pathak, Roth, and Sönmez (2005); Pathak and Sönmez (2008, 2013)).

[2]One-sided strategy-proofness holds only for agents with unit supply/demand (see Roth (1984a); Sönmez (1997)).

[3]Continuous transfers are present in real-world settings such as object assignment (Koopmans and Beckmann (1957); Shapley and Shubik (1971)), marriage (Becker (1973, 1974)), and labor market matching (Crawford and Knoer (1981); Kelso and Crawford (1982)).

[4]While many people believed or expected strategy-proofness results to hold for more general markets with continuous transfers, we are not aware that any such results were known formally prior to the work of Hatfield, Kojima, and Kominers (2017).

[5]On the other hand, the Vickrey–Clarke–Groves (VCG) mechanism is available in settings with continuously

Recently, Hatfield, Kojima, and Kominers (2017) proved one-sided strategy-proofness of deferred acceptance for unit-supply sellers in the trading network framework of Hatfield et al. (2013);[6,7] however, their proof strategy is quite indirect.[8] Here, we give the first *direct* proof of one-sided strategy-proofness in a context with continuously transferable utility. We consider worker–firm matching with quasilinear utility, and adapt the Hatfield and Milgrom (2005) method of proving one-sided strategy-proofness in settings with discrete contracts. Our approach makes use of a version of the Lone Wolf Theorem that we recently developed for settings with transferable utility (Jagadeesan et al. (2017)).

In recent independent work, Schlegel (2016) has shown that worker-optimal core-selecting mechanisms exist and are strategy-proof for workers in non-quasilinear settings that satisfy a preference substitutability assumption (see Theorem 5 of Schlegel (2016)). Schlegel's (2016) proof combines results on the strategy-proofness of worker-optimal core-selecting mechanisms in settings with discrete transfers (Hatfield and Milgrom (2005)) with a limiting argument. Our approach extends to Schlegel's (2016) setting by replacing our Lone Wolf Theorem (Jagadeesan et al. (2017)) with Schlegel's (2016) version of the lone wolf result.

The remainder of this paper is organized as follows: Section 2 presents an example that illustrates the idea and proof of our main result. Section 3 introduces a transferable-utility version of the Kelso and Crawford (1982) model of worker–firm matching. Section 4

transferable utility. Demange (1982), Leonard (1983), Demange et al. (1986), and Ausubel and Milgrom (2002) showed that the worker-optimal core-selecting mechanism yields VCG payoffs; they then used that fact to deduce one-sided strategy-proofness in one-to-one and single-seller settings. Our results show that the matching-theoretic approach to proving strategy-proofness generalizes to settings with transferable utility. In contrast, Ausubel (2004, 2006) and Sun and Yang (2014) work with multi-unit demand settings, where the matching-theoretic approach does not yield strategy-proofness results (Roth (1984a)).

[6]The Hatfield et al. (2013) model that Hatfield, Kojima, and Kominers (2017) work with is quite general—it embeds the worker–firm matching setting of Crawford and Knoer (1981) and the object allocation settings of Koopmans and Beckmann (1957), Shapley and Shubik (1971), Gul and Stacchetti (1999, 2000) and Sun and Yang (2006, 2009).

[7]Hatfield et al. (2013) assumed quasilinear utility (as we do), so their model does not fully generalize the model of Kelso and Crawford (1982); they (and we) only address the quasilinear case of Kelso and Crawford (1982).

[8]Effectively, Hatfield, Kojima, and Kominers (2017) showed that the seller-optimal stable matching mechanism coincides with VCG when all sellers have unit supply. We instead apply matching-theoretic arguments to prove strategy-proofness directly; by the Green–Laffont–Holmström Theorem, it then follows that the seller-optimal stable matching mechanism coincides with VCG.

introduces the Lone Wolf Theorem and proves our strategy-proofness result. Section 5 discusses antecedents and implications of our results.

2 An Illustrative Example

We begin with an example economy that illustrates our main result and proof strategy. We suppose that there are three workers (or as we might say, illustrators), Al, Bob, and Charles, and two firms, Gale Gallery and Shapley Studios. Firms' valuations are as given in Table 1; workers' costs of working are as presented in Table 2.

Workers employed	Gale Gallery's value	Shapley Studios's value
∅	0	0
{Al}	\$1000	\$1000
{Bob}	\$1000	\$800
{Charles}	\$800	\$600
{Al, Bob}	–	\$1600
{Al, Charles}	–	\$1400
{Bob, Charles}	–	\$1200
{Al, Bob, Charles}	–	–

Table 1: Firms' valuations for hiring different sets of employees in Section 2.

Employer	Al's cost	Bob's cost	Charles's cost
Gale Gallery	\$700	\$800	\$1000
Shapley Studios	\$300	\$300	\$300

Table 2: Workers' costs of working for each potential employer in Section 2.

We say that an employment outcome is in the *core* if no set of workers and firms can profitably recontract among themselves; such an outcome is *worker-optimal* if it is most-preferred by all workers among all core outcomes. Two worker-optimal core outcomes exist in our economy:[9]

[9]Note that both firms in our example have substitutable valuations (in the sense of Kelso and Crawford (1982)), so worker-optimal core outcomes are guaranteed to exist (see Kelso and Crawford (1982) and Hatfield et al. (2013)).

outcome $\mathcal{S}$, in which Al and Bob work for Shapley Studios at salaries of \$700 and \$500, respectively, and

outcome $\mathcal{T}$, in which Bob works for Gale Gallery at a salary of \$1000 and Al and Charles work for Shapley Studios at salaries of \$700 and \$300, respectively.

Note that, given their cost functions, all workers are indifferent between $\mathcal{S}$ and $\mathcal{T}$.

The Jagadeesan et al. (2017) Lone Wolf Theorem shows that if any agent is unmatched in one core outcome, then he receives 0 net utility in every core outcome.[10] For example, because Charles is unmatched in $\mathcal{S}$, he must receive 0 net utility in any core outcome.[11]

One mechanism for finding a worker-optimal core outcome is the *worker-proposing deferred acceptance mechanism (with salaries)*, also called the *descending salary adjustment process* (see Kelso and Crawford (1982)); under this mechanism, salaries start at their highest possible level and descend until all firms are willing to pay the salaries of the workers that prefer to work for them at those salaries, and the market clears. In our setting, if all workers reveal their true costs of employment (in their choices between firms at proposed salaries), then worker-proposing deferred acceptance selects either $\mathcal{S}$ or $\mathcal{T}$.

Our main result shows that worker-proposing deferred acceptance—indeed, any worker-optimal core-selecting mechanism—is strategy-proof for workers, that is, it is a dominant strategy for all workers to reveal their true costs/preferences. We explain the underlying logic in the case of Charles:[12] We can show by construction that if there is any strictly profitable deviation for Charles that assigns him to firm f at salary s, then there is a strictly profitable deviation in which Charles

1. represents that only firm f is acceptable and

2. correctly represents his cost of working at firm f.

[10]The Jagadeesan et al. (2017) Lone Wolf Theorem applies to competitive equilibria, but competitive equilibrium outcomes correspond to core outcomes (and to stable outcomes) in in many-to-one matching with transfers (Kelso and Crawford (1982); Roth (1984b); Hatfield et al. (2013)).

[11]In particular, observe that Charles receives 0 net utility under $\mathcal{T}$.

[12]The interested reader might consider possible deviation strategies to observe that none make Charles better-off.

As Charles is unemployed under $\mathcal{S}$, which is in the core when he reports his true preferences, $\mathcal{S}$ is clearly in the core when he misrepresents his preferences by shading his willingness to work. By the Lone Wolf Theorem, Charles must receive 0 net utility under any outcome in the core under his misrepresented preferences—but this contradicts our hypothesis that Charles has a strictly profitable deviation.[13]

3 Model

We work with a model of many-to-one matching with continuously transferable utility, following Crawford and Knoer (1981) and Kelso and Crawford (1982). There is a finite set F of *firms* and a finite set W of *workers*. We let $I \equiv F \cup W$ denote the full set of *agents*.

A firm f and a worker w can sign a *contract* (f, w, s) indicating that w will work for f at salary $s \in \mathbb{R}$. The full set of contracts is $X \equiv F \times W \times \mathbb{R}$. Given a set of contracts $A \subseteq X$, we denote the sets of contracts in A associated to firm $f \in F$ and worker $w \in W$ by

$$A_f \equiv A \cap (\{f\} \times W \times \mathbb{R}) = \{(f, w', s) : (f, w', s) \in A\} \text{ and}$$
$$A_w \equiv A \cap (F \times \{w\} \times \mathbb{R}) = \{(f', w, s) : (f', w, s) \in A\},$$

respectively. An *outcome* is a set of contracts $A \subseteq X$ under which each worker is employed by at most one firm, i.e., a set of contracts A for which $|A_w| \leq 1$ for all workers $w \in W$.

3.1 Preferences

Each worker $w \in W$ has a *valuation* over firms,

$$v_w : F \to \mathbb{R} \cup \{-\infty\};$$

[13] The fact that Charles receives 0 net utility under truthful reporting in our example simplifies the logic here. In the general argument, we first (again by construction) inflate costs of work to recover the 0-net-utility case.

as in our example, this valuation may encode the costs of producing products at each employer. The valuation v_w induces a quasilinear utility function $u_w : F \times \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$, defined by

$$u_w((f, s)) \equiv v_w(f) + s.$$

The utility function u_w naturally extends to outcomes $A \subseteq X$. Throughout, we use the convention that $u_w(\varnothing) = 0$.

Each firm f has a *valuation* function over sets of workers

$$v_f : \wp(W) \to \mathbb{R} \cup \{-\infty\},$$

normalized so that $v_f(\varnothing) = 0$.[14] The valuation v_f induces a quasilinear utility function $u_f : \wp(X_f) \to \mathbb{R} \cup \{-\infty\}$ defined by

$$u_f(A_f) \equiv \begin{cases} -\infty & |A_w| > 1 \text{ for some worker } w \\ v_f(\mathsf{w}(A_f)) - \sum_{(f,w,s) \in A_f} s & \text{otherwise,} \end{cases}$$

where $\mathsf{w}(Y)$ denotes the set of workers associated to contracts in Y. As with worker utility functions, the utility function u_f naturally extends to outcomes A. By construction, we have $u_f(\varnothing) = 0$.

3.2 The Core

An outcome A is in the *core (under the valuation profile v)* if there does not exist a *core block (for A)*, i.e., a set of agents $J \subseteq I$ and an outcome $B \subseteq X$ such that

- $B_i = \varnothing$ for all $i \notin J$;
- $u_j(B_j) > u_j(A_j)$ for all $j \in J$.

[14] Here, the notation $\wp$ denotes the *power set*, $\wp(R) \equiv \{R' : R' \subseteq R\}$ for any set R.

3.3 Mechanisms

The *set of possible valuations* for worker w is $V_w \equiv (\mathbb{R} \cup \{-\infty\})^F$. We let $V \equiv \bigtimes_{w \in W} V_w$ be the *set of possible valuation profiles*. A *mechanism* is a function $\mathcal{M} : V \to \wp(X)$ that maps valuation profiles to outcomes. A mechanism is *core-selecting* if it always returns core outcomes.

4 Strategy-Proofness of Worker-Optimal Mechanisms

A core-selecting mechanism $\mathcal{M}$ is *optimal for worker* w if for any core allocation A and any $v \in V$, we have

$$u_w(\mathcal{M}(v)) \geq u_w(A).$$

A mechanism $\mathcal{M}$ is *strategy-proof for worker* w if reporting truthfully is (weakly) dominant for worker w under $\mathcal{M}$—that is, if

$$u_w(\mathcal{M}(v)) \geq u_w\Big(\mathcal{M}\Big((v'_w, v_{W \smallsetminus \{w\}})\Big)\Big)$$

for all $v \in V$ and $v'_w \in V_w$.

Our main result is as follows.

Theorem 1. *Any core-selecting mechanism that is optimal for worker w is strategy-proof for w.*

4.1 Sketch of Proof

The key new ingredient in our proof of Theorem 1 is the following "Lone Wolf Theorem," which we have derived in other work.

Theorem 2 (Jagadeesan et al. (2017))**.** *If A and A' are both core outcomes, w is a worker, and we have $u_w(A_w) > 0$, then we must have $A'_w \neq \varnothing$. That is, if a worker receives strictly* *positive utility in some core outcome, then that worker is matched in every core outcome.*[15]

Theorem 2 is an analogue of the classical "lone wolf" theorem (see McVitie and Wilson (1970) and Gale and Sotomayor (1985), as well as Roth (1984a, 1986), Alkan (2002), Klaus and Klijn (2010), Hatfield and Kominers (2012), Hatfield and Milgrom (2005), Klijn and Yazıcı (2014), Ciupan et al. (2016) and Jagadeesan (2016)).[16] As Hatfield and Milgrom (2005) showed in the context of many-to-one matching with (discrete) contracts, the Lone Wolf Theorem is useful in proving one-sided strategy-proofness.

We adapt the Hatfield–Milgrom (2005) proof of (one-sided) strategy-proofness to deduce Theorem 1 from Theorem 2. Indeed, we suppose for the sake of deriving a contradiction that there exists a profitable deviation for worker w under core-selecting mechanism $\mathcal{M}$, namely reporting valuation $\bar{v}_w$ instead of v_w. By inflating w's costs of work sufficiently, we can assume that w is unmatched under truthful revelation but can obtain an individually rational match by deviating; from that individually rational match, we construct a core outcome under which w is matched, contradicting the Lone Wolf Theorem (Theorem 2).

5 Discussion

We have shown that any worker-optimal core-selecting mechanism is strategy-proof for workers in the context of matching with continuous transfers. We used a proof strategy based on the Lone Wolf Theorem. Due to the absence of a suitable Lone Wolf Theorem, it has not previously been clear whether such a strategy could be used in matching settings with continuous transfers, even though it has been applied in many other matching contexts.

[15]While Jagadeesan et al. (2017) formally state their Lone Wolf Theorem for competitive equilibria, Kelso and Crawford (1982) have shown that competitive equilibrium outcomes coincide with core outcomes in many-to-one matching markets with continuous transfers. Hence, the main result of Jagadeesan et al. (2017) directly implies Theorem 2.

[16]Recently, Schlegel (2016) has shown that Theorem 2 holds when workers' utility functions are continuous but not necessarily quasilinear, provided that all firms' valuations are *(grossly) substitutable* (in the sense of Kelso and Crawford (1982); see Footnote 19). Schlegel (2016) proved his result (Theorem 4 of Schlegel (2016)) by combining the Rural Hospitals Theorem for matching with contracts (Theorem 8 of Hatfield and Milgrom (2005)) with a limiting argument. As we have shown (Jagadeesan et al. (2017)), transferable utility substitutes for substitutability in Theorem 2, yielding a simpler proof.

5.1 Antecedents

Worker-optimal matching mechanisms have long been believed to be strategy-proof in many-to-one environments with transferable utility, but to our knowledge the only prior proof of this fact is a recent indirect argument due to Hatfield, Kojima, and Kominers (2017)—we give the first direct proof.[17]

Theorem 1 generalizes results of Demange (1982) and Leonard (1983) for one-to-one matching contexts (see also Demange et al. (1986) and Ausubel and Milgrom (2002)) and results of Gul and Stacchetti (2000) for object allocation settings. Our result is an analogue of Theorem 11 of Hatfield and Milgrom (2005), which gives the corresponding result for worker–firm matching with discrete (rather than continuous) transfers.[18] Our approach to Theorem 1 adapts the proof strategy of Hatfield and Milgrom (2005) to the context of matching with continuous transfers.

5.2 Implications

Theorem 4 of Hatfield et al. (2013) implies the existence of worker-optimal competitive equilibria if all firms have substitutable valuations.[19] Moreover, Kelso and Crawford (1982) showed that there exist competitive equilibrium prices that support any core allocation. Thus, *a worker-optimal core-selecting mechanism exists—and hence, by our results, strategy-proof worker–firm matching is possible—if all firms have substitutable valuations.* In particular, our results show that the worker-proposing deferred acceptance mechanism (or equivalently,

[17]Hatfield, Kojima, and Kominers (2017) proved Theorem 1 by showing that worker-optimal mechanisms are efficient and guarantee each worker the full marginal surplus from a change in valuation; these observations together imply strategy-proofness by the main result of Hatfield, Kojima, and Kominers (2017).

[18]As we discussed in Section 1, analogues of Theorem 1 have been found in a range of matching contexts without transferable utility, including marriage and college admissions matching (Dubins and Freedman (1981); Roth (1982); Martínez et al. (2004)), many-to-one matching with contracts (Hatfield and Milgrom (2005); Hatfield and Kojima (2009, 2010); Hatfield and Kominers (2015); Hatfield, Kominers, and Westkamp (2016)), and supply chain matching (Hatfield and Kominers (2012); see also Ostrovsky (2008)).

[19]The valuation of firm f is *(grossly) substitutable* if the corresponding indirect utility function is submodular, or equivalently if an increase in the salary f must pay some worker cannot reduce f's demand for workers whose salaries are unchanged (see, e.g., Kelso and Crawford (1982); Ausubel and Milgrom (2002); Hatfield et al. (2017)).

the worker-optimal core-selecting matching mechanism) is strategy-proof for workers in the Kelso–Crawford (1982) setting with continuous transfers and quasilinear utility.

5.3 Matching with Contracts

Our proof of Theorem 1 extends *verbatim* to settings in which workers and firms also negotiate over non-pecuniary contract terms (Roth, 1984b; Hatfield and Milgrom, 2005; Hatfield et al., 2013). However, we need to maintain the assumption that all agents' preferences over transfers are quasilinear.

5.4 Group Strategy-Proofness

Demange and Gale (1985) and Demange (1987) proved group strategy-proofness for one-to-one matching with continuous transfers.[20] However, no group strategy-proofness result is known in our setting.

Group strategy-proofness results have been found in nearly all the matching contexts with discrete contracts that have individual strategy-proofness results (see Section 5.1 and Footnote 18 for references), via an argument first introduced by Hatfield and Kojima (2009). Unfortunately, the Hatfield and Kojima (2009) approach cannot be used directly in settings with continuous transfers.[21] It seems likely, however, that the techniques of Schlegel (2016) can show that worker-optimal core-selecting mechanisms are group strategy-proof in settings with continuous transfers—at least when all firms' valuations are substitutable and quasilinear.

[20] A mechanism $\mathcal{M}$ is *(weakly) group strategy-proof (for workers)* if for all $S \subseteq W$, $v \in V$, and $v' \in \bigtimes_{w \in S} V_w$, we must have

$$u_w(\mathcal{M}(v)) \geq u_w(\mathcal{M}((v', v_{W \smallsetminus S})))$$

for some $w \in S$—that is, there is no deviation $v' \in \bigtimes_{w \in S} V_w$ that makes all workers in S strictly better off than under truthful reporting.

[21] The technical difficulty is that the preference modifications used by Hatfield and Kojima (2009) generate discontinuous income effects.

A Proof of Theorem 1

For the remainder of this Appendix, we fix a worker $w \in W$. Throughout, we use the convention that utility functions u_w, $\bar{u}_w$ and $\hat{u}_w$ are associated to valuations v_w, $\bar{v}_w$, and $\hat{v}_w$, respectively, via the utility function construction established in Section 3.

A.1 Preliminaries

The proof of Theorem 1 makes use of two lemmata. Lemma A.1 asserts that any core outcome remains in the core if firms become weakly less desirable to the workers that are not their employees. Lemma A.2 asserts that increasing the utility of the outside option only shrinks the core by making certain outcomes fail to be individually rational.

Lemma A.1. *Let $\bar{v}_w \in V_w$ be a valuation with $\bar{v}_w \leq v_w$ and let A be an outcome with $u_w(A_w) = \bar{u}_w(A_w)$. If A be a core outcome under the valuation profile v, then A is a core outcome under the valuation profile $\left(\bar{v}_w, v_{W \smallsetminus \{w\}}\right)$.*

Proof. We prove the contrapositive. Assume that A is an outcome with $u_w(A_w) = \bar{u}_w(A_w)$ and that coalition J and outcome B together constitute a core block for A under the valuation profile $\left(\bar{v}_w, v_{W \smallsetminus \{w\}}\right)$.

If $w \notin J$, then J and B clearly constitute a core block of A for the valuation profile v, so that A is not a core outcome for the valuation profile v. Thus, we suppose that $w \in J$. As $\bar{v}_w \leq v_w$ by assumption, we have

$$u_w(B_w) \geq \bar{u}_w(B_w) > \bar{u}_w(A_w) = u_w(A_w). \tag{1}$$

It follows from (1) that the coalition J and the outcome B constitute a core block for A under the valuation profile v. Thus, we see that A is not a core outcome under the valuation profile v, as desired. □

Lemma A.2. *Consider an outcome A and let $\varepsilon \geq 0$. We let $\bar{v}_w = v_w - \varepsilon$. If A is a core outcome under the valuation profile $\left(\bar{v}_w, v_{W \smallsetminus \{w\}}\right)$ and $A_w \neq \varnothing$, then A is a core outcome under the valuation profile v.*

Proof. We prove the contrapositive. Assume that $A_w \neq \varnothing$ and that coalition J and outcome B constitute a core block for A under the valuation profile v. We can assume that $w \in J$, as otherwise J and B clearly constitute a core block for A under the valuation profile $\left(v_w - \varepsilon, v_{W \smallsetminus \{w\}}\right) = \left(\bar{v}_w, v_{W \smallsetminus \{w\}}\right)$.

Supposing that $w \in J$, we have

$$\bar{u}_w(B) \geq u_w(B_w) - \varepsilon > u_w(A_w) - \varepsilon = \bar{u}_w(A), \tag{2}$$

where the the second inequality in (2) holds because J and B constitute a core block for A under valuation profile v, and the equality in (2) holds because $A_w \neq \varnothing$.[22]

It follows from (2) that the coalition J and the outcome B constitute a core block for A under the valuation profile $\left(v_w - \varepsilon, v_{W \smallsetminus \{w\}}\right) = \left(\bar{v}_w, v_{W \smallsetminus \{w\}}\right)$. Thus, we find that A is not a core outcome under the valuation profile $\left(\bar{v}_w, v_{W \smallsetminus \{w\}}\right)$, as desired. ☐

A.2 Main Argument

We let $\mathcal{M}$ be a core-selecting mechanism that is optimal for w. We let $v \in V$ and $\bar{v}_w \in V_w$ be arbitrary, and take $A = \mathcal{M}(v)$ and $A' = \mathcal{M}\left((\bar{v}_w, v_{W \smallsetminus \{w\}})\right)$. We suppose for the sake of deriving a contradiction that $u_w(A_w) < u_w(A'_w)$.

As $u_w(A'_w) > u_w(A_w) \geq 0$, worker w receives strictly positive utility under outcome A'. Thus, w is matched to a firm under outcome A'; we denote this firm by f_w. We let $\eta > 0$ be such that

$$u_w(A_w) < \eta < u_w(A'_w). \tag{3}$$

[22]To see the first inequality in (2), we note that by our choice of $\bar{v}$, we have $\bar{u}_w(B) = u_w(B_w) - \varepsilon$ except when $\bar{u}_w(B) = 0$, in which case $\bar{u}_w(B) = 0 > -\varepsilon = u_w(B_w) - \varepsilon$.

We let $A'' = \mathcal{M}\big((v_w - \eta, v_{W \smallsetminus \{w\}})\big)$.

Claim. *We have* $A''_w = \varnothing$.

Proof. Suppose for the sake of deriving a contradiction that $A''_w \neq \varnothing$. Because $\mathcal{M}$ is core-selecting, Lemma A.2 then implies that A'' is a core outcome with respect to the valuation profile v. The individual rationality of A'' with respect to $\big(v_w - \eta, v_{W \smallsetminus \{w\}}\big)$ implies that $u_w(A''_w) > \eta$, which contradicts the worker-optimality of $\mathcal{M}$, as $\eta > u_w(A_w)$ (recall (3)). $\square$

Now, we define valuation $\hat{v} \in V_w$ by

$$\hat{v}_w(f) = \begin{cases} v_w(f) - \eta & f = f_w \\ -\infty & \text{otherwise;} \end{cases}$$

effectively, under $\hat{v}_w$, w values only f_w—and values f_w at an amount strictly bounded above by $v_w(f_w)$. By construction, we have

$$\hat{v}_w \leq v_w - \eta. \tag{4}$$

We let $\bar{\eta} = \bar{v}_w(f_w) - v_w(f_w) + \eta$. With (4) and our choice of $\bar{\eta}$, we have

$$\hat{v}_w \leq v_w - \eta = \bar{v}_w - \bar{\eta}; \tag{5}$$

moreover, we have

$$\hat{u}_w(A') = \bar{u}_w(A') - \bar{\eta} \tag{6}$$

$$= u_w(A') - \eta > 0. \tag{7}$$

Now, as A' is a core outcome under valuation profile $(\bar{v}_w, v_{W \smallsetminus \{w\}})$ and w receives strictly positive utility at A' under the valuation $\bar{v}_w - \bar{\eta}$, we know from our choice of η that A' must also be a core outcome under the valuation profile $\big(\bar{v}_w - \bar{\eta}, v_{W \smallsetminus \{w\}}\big)$. By Lemma A.1, it then follows from (5) and (6) that A' is a core outcome under the valuation profile $(\hat{v}_w, v_{W\smallsetminus\{w\}})$. As $A''_w = \varnothing$, combining Lemma A.1 and (5) implies that A'' is also a core outcome under the valuation profile $(\hat{v}_w, v_{W\smallsetminus\{w\}})$.[23]

But then, we have both A' and A'' in the core under the valuation profile $(\hat{v}_w, v_{W\smallsetminus\{w\}})$. As $u''_w(A'_w) > 0$ by (7) and $A''_w = \varnothing$, this contradicts the Lone Wolf Theorem (Theorem 2).

References

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005). The New York City high school match. *American Economic Review 95*(2), 364–367.

Abdulkadiroğlu, A., P. A. Pathak, A. E. Roth, and T. Sönmez (2005). The Boston public school match. *American Economic Review 95*(2), 368–371.

Abdulkadiroğlu, A. and T. Sönmez (2003). School choice: A mechanism design approach. *American Economic Review 93*(3), 729–747.

Alkan, A. (2002). A class of multipartner matching markets with a strong lattice structure. *Economic Theory 19*(4), 737–746.

Ausubel, L. M. (2004). An efficient ascending-bid auction for multiple objects. *American Economic Review 94*(5), 1452–1475.

Ausubel, L. M. (2006). An efficient dynamic auction for heterogeneous commodities. *American Economic Review 96*(3), 602–629.

Ausubel, L. M. and P. R. Milgrom (2002). Ascending auctions with package bidding. *Advances in Theoretical Economics 1*(1).

[23] We have $\hat{u}_w(A''_w) = \hat{u}_w(\varnothing) = 0 = \check{u}_w(\varnothing) = \check{u}_w(A''_w)$, where

$$\check{u}_w(B) = \begin{cases} \bar{v}_w(B) - \bar{\eta} & B \neq \varnothing \\ 0 & B = \varnothing \end{cases}$$

is the utility function associated to the valuation function $\bar{v}_w - \bar{\eta}$.

Balinski, M. and T. Sönmez (1999). A tale of two mechanisms: Student placement. *Journal of Economic Theory 84*(1), 73–94.

Becker, G. S. (1973). A theory of marriage: Part I. *Journal of Political Economy 81*(1), 813–846.

Becker, G. S. (1974). A theory of marriage: Part II. *Journal of Political Economy 82*(2), 11–26.

Ciupan, A., J. W. Hatfield, and S. D. Kominers (2016). An elementary proof of the lone wolf theorem. Working paper.

Clarke, E. H. (1971). Multipart pricing of public goods. *Public Choice 11*(1), 17–33.

Crawford, V. P. and E. M. Knoer (1981). Job matching with heterogeneous firms and workers. *Econometrica 49*(2), 437–450.

Demange, G. (1982). Strategyproofness in the assignment market game. Working paper.

Demange, G. (1987). Nonmanipulable cores. *Econometrica 55*(5), 1057–1074.

Demange, G. and D. Gale (1985). The strategy structure of two-sided matching markets. *Econometrica 53*(4), 873–888.

Demange, G., D. Gale, and M. Sotomayor (1986). Multi-item auctions. *Journal of Political Economy 94*(4), 863–872.

Dubins, L. E. and D. A. Freedman (1981). Machiavelli and the Gale-Shapley algorithm. *American Mathematical Monthly 88*(7), 485–494.

Gale, D. and L. S. Shapley (1962). College admissions and the stability of marriage. *American Mathematical Monthly 69*(1), 9–15.

Gale, D. and M. Sotomayor (1985). Ms. Machiavelli and the stable matching problem. *American Mathematical Monthly 92*(4), 261–268.

Green, J. and J.-J. Laffont (1977). Characterization of satisfactory mechanisms for the revelation of preferences for public goods. *Econometrica 45*(2), 427–438.

Groves, T. (1973). Incentives in teams. *Econometrica 41*(4), 617–631.

Gul, F. and E. Stacchetti (1999). Walrasian equilibrium with gross substitutes. *Journal of Economic Theory 87*(1), 95–124.

Gul, F. and E. Stacchetti (2000). The English auction with differentiated commodities. *Journal of Economic Theory 92*(1), 66–95.

Hatfield, J. W. and F. Kojima (2009). Group incentive compatibility for matching with contracts. *Games and Economic Behavior 67*(2), 745–749.

Hatfield, J. W. and F. Kojima (2010). Substitutes and stability for matching with contracts. *Journal of Economic Theory 145*(5), 1704–1723.

Hatfield, J. W., F. Kojima, and S. D. Kominers (2017). Strategy-proofness, investment efficiency, and marginal returns: An equivalence. Becker Friedman Institute for Research in Economics Working Paper.

Hatfield, J. W. and S. D. Kominers (2012). Matching in networks with bilateral contracts. *American Economic Journal: Microeconomics 4*(1), 176–208.

Hatfield, J. W. and S. D. Kominers (2015). Hidden substitutes. Working paper.

Hatfield, J. W., S. D. Kominers, A. Nichifor, M. Ostrovsky, and A. Westkamp (2013). Stability and competitive equilibrium in trading networks. *Journal of Political Economy 121*(5), 966–1005.

Hatfield, J. W., S. D. Kominers, A. Nichifor, M. Ostrovsky, and A. Westkamp (2017). Full substitutability. Working paper.

Hatfield, J. W., S. D. Kominers, and A. Westkamp (2016). Stability, strategy-proofness, and cumulative offer mechanisms. Working paper.

Hatfield, J. W. and P. R. Milgrom (2005). Matching with contracts. *American Economic Review 95*(4), 913–935.

Holmström, B. (1979). Groves' scheme on restricted domains. *Econometrica 47*(5), 1137–1144.

Jagadeesan, R. (2016). Lattice structure and the rural hospitals theorem in matching with complementarities. In preparation.

Jagadeesan, R., S. D. Kominers, and R. Rheingans-Yoo (2017). Lone wolves in competitive equilibria. Working paper.

Kelso, A. S. and V. P. Crawford (1982). Job matching, coalition formation, and gross substitutes. *Econometrica 50*(6), 1483–1504.

Klaus, B. and F. Klijn (2010). Smith and Rawls share a room: Stability and medians. *Social Choice and Welfare 35*(4), 647–667.

Klijn, F. and A. Yazıcı (2014). A many-to-many 'rural hospital theorem'. *Journal of Mathematical Economics 54*, 63–73.

Koopmans, T. C. and M. Beckmann (1957). Assignment problems and the location of economic activities. *Econometrica 25*(1), 53–76.

Leonard, H. B. (1983). Elicitation of honest preferences for the assignment of individuals to positions. *Journal of Political Economy 91*(3), 461–479.

Martínez, R., J. Massó, A. Neme, and J. Oviedo (2004). On group strategy-proof mechanisms for a many-to-one matching model. *International Journal of Game Theory 33*(1), 115–128.

McVitie, D. G. and L. B. Wilson (1970). Stable marriage assignment for unequal sets. *BIT Numerical Mathematics 10*(3), 295–309.

Ostrovsky, M. (2008). Stability in supply chain networks. *American Economic Review 98*(3), 897–923.

Pathak, P. A. and T. Sönmez (2008). Leveling the playing field: Sincere and sophisticated players in the Boston mechanism. *American Economic Review 98*(4), 1636–1652.

Pathak, P. A. and T. Sönmez (2013). School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation. *American Economic Review 103*(1), 80–106.

Roth, A. E. (1982). The economics of matching: Stability and incentives. *Mathematics of Operations Research 7*(4), 617–628.

Roth, A. E. (1984a). The evolution of the labor market for medical interns and residents: A case study in game theory. *Journal of Political Economy 92*(6), 991–1016.

Roth, A. E. (1984b). Stability and polarization of interests in job matching. *Econometrica 52*(1), 47–57.

Roth, A. E. (1986). On the allocation of residents to rural hospitals: A general property of two-sided matching markets. *Econometrica 54*(2), 425–427.

Roth, A. E. and E. Peranson (1999). The redesign of the matching market for American physicians: Some engineering aspects of economic design. *American Economic Review 89*(4), 748–780.

Roth, A. E. and M. A. O. Sotomayor (1990). *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press.

Schlegel, J. C. (2016). Virtual demand and stable mechanisms. Working paper.

Shapley, L. S. and M. Shubik (1971). The assignment game I: The core. *International Journal of Game Theory 1*(1), 111–130.

Sönmez, T. (1997). Manipulation via capacities in two-sided matching markets. *Journal of Economic Theory 77*(1), 197–204.

Sun, N. and Z. Yang (2006). Equilibria and indivisibilities: Gross substitutes and complements. *Econometrica 74*(5), 1385–1402.

Sun, N. and Z. Yang (2009). A double-track adjustment process for discrete markets with substitutes and complements. *Econometrica 77*(3), 933–952.

Sun, N. and Z. Yang (2014). An efficient and incentive compatible dynamic auction for multiple complements. *Journal of Political Economy 122*(2), 422–466.

Vickrey, W. (1961). Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance 16*(1), 8–37.