Lone Wolves in Competitive Equilibria*

Ravi Jagadeesan
Harvard Business School and Department of Economics
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
September 29, 2019

Abstract

This paper develops a class of equilibrium-independent predictions of competitive equilibrium with indivisibilities. Specifically, we prove an analogue of the “Lone Wolf Theorem” of classical matching theory for the Baldwin and Klemperer (2019) model of exchange economies with transferable utility, showing that any agent who does not participate in trade in one competitive equilibrium must receive her autarky payoff in every competitive equilibrium. Our results extend to approximate equilibria and to settings in which utility is only approximately transferable.

JEL Classification: C78, D47, D51

Keywords: Indivisibilities; Matching; Lone Wolf Theorem

---

*This paper was previously circulated under the more amusing title “If You’re Happy and You Know It Then You’re Matched.” 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 Mohammad Akbarpour, John William Hatfield, Bettina Klaus, Shengwu Li, Paul Milgrom, Robbie Minton, Casey Mulligan, Alvin E. Roth, Alexander Teytelboym, the editor, Marc Fleurbaey, and the anonymous referee for helpful comments. Jagadeesan gratefully acknowledges the support of a National Science Foundation Graduate Research Fellowship under grant DGE-1745303, as well as travel support from Harvard Business School and the Harvard Mathematics Department. Kominers gratefully acknowledges the support of National Science Foundation grant SES-1459912 and the Ng Fund and the Mathematics in Economics Research Fund of the CMSA.
1 Introduction

In models of exchange and production, results on the uniqueness of equilibria typically require continuity and convexity conditions on agents’ preferences, in addition to further conditions on aggregate demand (see, for example, Section 17.F of Mas-Colell et al. (1995)). However, continuity and convexity conditions fail in the presence of indivisibilities, leading to the existence of multiple equilibria. In turn, the presence of multiple equilibria weakens the predictive power of economic models.

To avoid the problems of multiplicity, we often consider predictions that do not depend on which equilibrium is realized. In this paper, we develop a class of equilibrium-independent predictions of competitive equilibrium in the Baldwin and Klemperer (2019) model of exchange economies with indivisible goods and transferable utility. Specifically, we show that any agent who does not participate in trade in one competitive equilibrium must receive her autarky payoff in every competitive equilibrium. Hence, observing that an agent does not participate in trade indicates that there is no competitive equilibrium in which that agent improves on her autarky payoff.

Our result is analogous to the classical “Lone Wolf Theorem” from matching theory, which asserts that in one-to-one matching without transfers, any agent who is unmatched in some stable outcome is unmatched in every stable outcome (McVitie and Wilson (1970)). Informally, the Lone Wolf Theorem shows the set of “lone wolves” (unmatched agents) is equilibrium-independent. Since our setting involves continuous transfers, and hence indifferences, we cannot obtain sharp predictions regarding the set of agents that participate in

\[1\] In structural estimation, for example, it is possible to avoid the problem of multiplicity of equilibria by performing inference based on equilibrium-independent predictions of the model (see, for example, Bresnahan and Reiss (1991)). A related alternative approach, which is valid even when there are few equilibrium-independent predictions, is to obtain set-identification of the parameters of interest by assuming that the observation is one of many possible equilibria (see, for example, Tamer (2003); Ciliberto and Tamer (2009); Pakes (2010); Galichon and Henry (2011); Pakes et al. (2015); Bontemps and Magnac (2017)).


\[3\] To the best of our knowledge, the term “Lone Wolf Theorem” was coined by Klaus and Klijn (2010).
trade in equilibrium. Instead, we show that any agent who does not participate in trade in one equilibrium cannot benefit from trade in any equilibrium.

When utility is perfectly transferable, there is generically a unique efficient allocation of goods among agents, although there may be multiple possible equilibrium price vectors. Hence, the set of agents that participate in trade is generically equilibrium-independent. However, our lone wolf result applies even when there are multiple efficient allocations. Moreover, we show that versions of our Lone Wolf Theorem hold in settings in which multiple allocations can be robustly sustained in equilibrium. Specifically, we consider an approximate equilibrium concept—which may be a reasonable solution when the economy suffers from small frictions—and derive an approximate Lone Wolf Theorem. We also prove an approximate Lone Wolf Theorem for equilibria in economies in which utility is only approximately transferable among agents.

Quasi-conversely, we show that some assumption on the transferability of utility is essential to our lone wolf results. Specifically, we show by example that lone wolf results do not generally hold in settings with strong income effects. Hence, while our lone wolf results show that competitive equilibrium has some (approximately) robust predictions when utility is (approximately) transferable, it remains an open question whether we can derive classes of equilibrium-independent predictions more generally.

In a companion paper (Jagadeesan et al. (2018)), we use our Lone Wolf Theorem to show that descending salary-adjustment processes are strategy-proof in quasilinear Kelso and Crawford (1982) economies. Our argument for strategy-proofness rests crucially on applying our Lone Wolf Theorem to (non-generic) economies with multiple efficient matchings.

The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 presents our lone wolf results. Section 4 discusses the role of the hypothesis that utility is (approximately) transferable. Section 5 connects our results to the matching literature. Proofs omitted from the main text are presented in Appendix A.
2 Model

We consider a general model of exchange economies with indivisible goods and transferable utility, following Baldwin and Klemperer (2019); this framework embeds the trading network model of Hatfield et al. (2013), which in turn generalizes the settings of Gul and Stacchetti (1999, 2000) and Sun and Yang (2006, 2009). There is a finite set $\Gamma$ of goods and a finite set $I$ of agents.

Each agent has an endowment $e^i \in \mathbb{Z}^\Gamma$ and a valuation

$$v^i : \mathbb{Z}^\Gamma \to \mathbb{R} \cup \{-\infty\}.$$  

We assume that valuations are bounded above and that $v^i(e^i) > -\infty$; however, we do not impose any further conditions on them. The valuation $v^i$ induces a \textit{quasilinear} utility function $u^i$ by

$$u^i(q^i, t^i) = v^i(q^i) + t^i,$$

where $t^i$ is the transfer received by agent $i$. We allow $t^i$ to be arbitrarily negative, so agents have unlimited budgets.

A competitive equilibrium is comprised of (1) an allocation of goods to agents and (2) prices for each good such that the allocation maximizes each agent’s utility given prices.

**Definition 1.** A \textit{competitive equilibrium} is a pair $[q; p]$, where $q = (q^i)_{i \in I}$, at which

- each agent $i$ demands her bundle $q^i$ at prevailing prices $p$, i.e.,

$$q^i \in \arg \max_{\bar{q}^i \in \mathbb{Z}^\Gamma} \big\{ u^i(\bar{q}^i, (e^i - \bar{q}^i) \cdot p) \big\}$$  

for all $i \in I$, and

- the market clears, i.e.,

$$\sum_{i \in I} q^i = \sum_{i \in I} e^i.$$
Baldwin and Klemperer (2019) have provided sufficient conditions for the existence of competitive equilibria in the model we consider here. For example, competitive equilibria exist if units of goods are substitutable and only finitely many bundles have valuations greater than $-\infty$ (see also Kelso and Crawford (1982), Gul and Stacchetti (1999), and Milgrom and Strulovici (2009)).

3 Results

The classical Lone Wolf Theorem from the theory of two-sided matching without transfers asserts that any agent who is matched in some equilibrium outcome is matched in every equilibrium (McVitie and Wilson (1970); see also Roth (1984, 1986) and Klaus and Klijn (2010)). In this section, we prove analogues of the Lone Wolf Theorem for exchange economies.

3.1 An Exact Lone Wolf Theorem

For a pair $[q; p]$, we say that an agent $i \in I$ does not participate in trade in $[q; p]$ if $q^i = e^i$. Such an agent is a “lone wolf” in $[q; p]$.

Our first result asserts that any agent who does not participate in trade in some competitive equilibrium receives her autarky payoff in every competitive equilibrium. Thus, we show that any agent who is a “lone wolf” in some equilibrium cannot benefit from trade in any equilibrium.

**Theorem 1.** Let $[q; p]$ and $[\hat{q}; \hat{p}]$ be competitive equilibria, and let $j \in I$ be an agent. If $j$ does not participate in trade in $[\hat{q}; \hat{p}]$ (i.e., if $\hat{q}^j = e^j$), then $u^j(q^j, t^j) = u^j(e^j, 0)$, where $t^j = (e^j - q^j) \cdot p$ is the net transfer to agent $j$ at equilibrium $[q; p]$.

We prove Theorem 1 as a corollary of a more general result that allows for optimization error in equilibrium.
3.2 An Approximate Lone Wolf Theorem

Note that Theorem 1 has bite only in economies in which there are multiple competitive equilibrium allocations. As competitive equilibrium allocations are generically unique in transferable utility economies, Theorem 1 has nontrivial consequences only for nongeneric economies.

However, as we show in this section, our lone wolf result holds more generally. In particular, we prove a version of Theorem 1 for approximate equilibria; as approximate equilibria are only approximately efficient, this result has nontrivial consequences in settings in which equilibrium allocations are robustly non-unique.

We relax the definition of competitive equilibrium by allowing agents’ total maximization error to be positive but bounded above by \( \varepsilon \).

**Definition 2.** An \( \varepsilon \)-equilibrium consists of a pair \( [q; p] \) for which

\[
\sum_{i \in I} \left( \max_{q^i \in \mathbb{Z}^n} \left\{ u^i(q^i, (e^i - q^i) \cdot p) \right\} - u^i(q^i, (e^i - q^i) \cdot p) \right) \leq \varepsilon
\]

and (2) is satisfied (i.e., the market clears).

Our Approximate Lone Wolf Theorem asserts that if there exists an \( \varepsilon \)-equilibrium in which no agents in \( J \subseteq I \) participate in trade, then the difference between the total utility of agents in \( J \) and the total autarky payoff of agents in \( J \) is bounded above by \( (\delta + \varepsilon) \) in every \( \delta \)-equilibrium.

**Theorem 2.** Let \( [q; p] \) be a \( \delta \)-equilibrium, let \( [\hat{q}; \hat{p}] \) be an \( \varepsilon \)-equilibrium, and let \( J \subseteq I \) be a set of agents. If no agent in \( J \) participates in trade in \( [\hat{q}; \hat{p}] \) (i.e., if \( \hat{q}^j = e^j \) for all \( j \in J \)), then

\[
\sum_{j \in J} u^j(q^j, t^j) - \sum_{j \in J} u^j(e^j, 0) \leq \delta + \varepsilon,
\]

where \( t^j = (e^j - q^j) \cdot p \) is the net transfer to agent \( j \) at approximate equilibrium \( [q; p] \).
The key to the proof of Theorem 2 is the following lemma, which allows us to produce a new approximate equilibrium from any two approximate equilibria.

**Lemma 1.** If \([q; p]\) is a \(\delta\)-equilibrium and \([\hat{q}; \hat{p}]\) is an \(\varepsilon\)-equilibrium, then \([\hat{q}; p]\) is a \((\delta + \varepsilon)\)-equilibrium.

Lemma 1, which we prove in Appendix A, is an approximate version of the well-known fact that \([\hat{q}; p]\) is a competitive equilibrium whenever \([q; p]\) and \([\hat{q}; \hat{p}]\) are competitive equilibria (see page 3 of Shapley (1964), as well as Gul and Stacchetti (1999), Sun and Yang (2006), and Hatfield et al. (2013)).

To prove Theorem 2, we exploit the fact that \([\hat{q}; p]\) is a \((\delta + \varepsilon)\)-equilibrium (by Lemma 1), so the allocation \(\hat{q}\) must approximately maximize \(j\)'s utility given the price vector \(p\).

**Proof of Theorem 2.** By Lemma 1, \([\hat{q}; p]\) is a \((\delta + \varepsilon)\)-equilibrium. As \(\hat{q}^j = e^j\) for all \(j \in J\) by assumption, we have that

\[
\sum_{j \in J} \left( \max_{\bar{q}^j \in \mathbb{Z}^k} \{ u^j(\bar{q}^j, (e^j - \bar{q}^j) \cdot p) \} - u^j(e^j, 0) \right) \leq \delta + \varepsilon. \tag{3}
\]

Meanwhile, with \(t^j = (e^j - q^j) \cdot p\), we have that

\[
u^j(q^j, t^j) = u^j(q^j, (e^j - q^j) \cdot p) \leq \max_{\bar{q}^j \in \mathbb{Z}^k} \{ u^j(\bar{q}^j, (e^j - \bar{q}^j) \cdot p) \}. \tag{4}\]

Combining (3) and (4), we see that

\[
\sum_{j \in J} \left( u^j(q^j, t^j) - u^j(e^j, 0) \right) \leq \sum_{j \in J} \left( \max_{\bar{q}^j \in \mathbb{Z}^k} u^j(\bar{q}^j, (e^j - \bar{q}^j) \cdot p) - u^j(e^j, 0) \right) \leq \delta + \varepsilon,
\]

as desired.

Taking \(J = \{j\}\) yields the following corollary.
Corollary 1. Let \([q; p]\) be a \(\delta\)-equilibrium, let \([\hat{q}; \hat{p}]\) be an \(\varepsilon\)-equilibrium, and let \(j \in I\) be an agent. If \(j\) does not participate in trade in \([\hat{q}; \hat{p}]\) (i.e., if \(\hat{q}^j = e^j\)), then

\[
W^j(q^j, t^j) - W^j(e^j, 0) \leq \delta + \varepsilon,
\]

where \(t^j = (e^j - q^j) \cdot p\).

Note that the \(\delta = \varepsilon = 0\) case of Corollary 1 is Theorem 1.

3.3 A Lone Wolf Theorem for Economies with Approximately Transferable Utility

Theorem 2 also implies a Lone Wolf Theorem for economies in which utility is “close to transferable” in a formal sense. Specifically, we consider economies in which agents’ utility functions can be well-approximated by quasilinear utility functions.

Definition 3. A utility function \(u^i\) is quasilinear within \(\eta\) if there exists a quasilinear utility function \(\tilde{u}^i\) for which

\[
\sup_{q^i, t^i} \{|u^i(q^i, t^i) - \tilde{u}^i(q^i; t^i)|\} \leq \eta.
\]

When utility functions are approximately quasilinear, we obtain another approximate Lone Wolf Theorem.\(^4\)

Theorem 3. Let \([q; p]\) and \([\hat{q}; \hat{p}]\) be competitive equilibria, and let \(J \subseteq I\) be a set of agents. If no agent in \(J\) participates in trade in \([\hat{q}; \hat{p}]\) (i.e., if \(\hat{q}^j = e^j\) for all \(j \in J\)) and \(u^i\) is quasilinear within \(\eta\) for all \(i \in I\), then

\[
\sum_{j \in J} W^j(q^j, t^j) - \sum_{j \in J} W^j(e^j, 0) \leq 6\eta|I|,
\]

\(^4\)While Definition 1 technically assumes that utility functions are quasilinear, the definition extends \textit{verbatim} to settings with income effects. Hence, we use the notion of competitive equilibrium from Definition 1 in this section.
where \( t^j = (e^j - q^j) \cdot p \) is the net transfer to agent \( j \) at equilibrium \([q; p]\).

To prove Theorem 3, we construct an approximating economy in which utility functions are quasilinear. Competitive equilibria in the original economy give rise to approximate equilibria in the approximating economy; Theorem 3 then follows from Theorem 2.

**Proof of Theorem 3.** We define an auxiliary economy in which utility functions are given by \((\hat{u}^i)_{i \in I}\), with each \( \hat{u}^i \) quasilinear as in Definition 3. By construction, \([q; p] \) and \([\hat{q}; \hat{p}] \) are \( 2\eta|I| \)-equilibria in the auxiliary economy. Theorem 2 then guarantees that

\[
\sum_{j \in J} \hat{u}^j(q^j; t^j) - \sum_{j \in J} \hat{u}^j(e^j; 0) \leq 4\eta|I|.
\] (5)

As \( \sum_{j \in J} |u^j(q^j; t^j) - \hat{u}^j(q^j; t^j)| \leq \eta|I| \) and \( \sum_{j \in J} |u^j(e^j; 0) - \hat{u}^j(e^j; 0)| \leq \eta|I| \) by our choice of \((\hat{u}^i)_{i \in I}\), we see from (5) that

\[
\sum_{j \in J} u^j(q^j; t^j) - \sum_{j \in J} u^j(e^j, 0) \leq 6\eta|I|,
\]

as desired. \( \square \)

### 4 The Role of Utility Transferability

The proofs of our Lone Wolf Theorems use the assumption that utility is (at least approximately) transferable. In fact, the Lone Wolf Theorem does not always hold when utility is not transferable. For example, the result can fail in the presence of income effects.

Consider an economy with four agents, a house builder \( B \), two real estate agents \( \text{HighEnd} \) and \( \text{LowEnd} \), and a consumer \( C \). Real estate agent \( \text{HighEnd} \) specializes in high-end properties, while \( \text{LowEnd} \) specializes in low-end properties. The builder \( B \) can construct a high-end property or a low-end property and sell it to the consumer \( C \) via \( \text{HighEnd} \) or \( \text{LowEnd} \), respectively.
These possible interactions can be summarized in the following network.\(^5\)

\[
\begin{array}{c}
\text{B} \\
\text{HighEnd} & \text{LowEnd} \\
\text{C}
\end{array}
\]

\(^6\)Formally, the trading network is a case of the model of Fleiner et al. (2019), who incorporate income effects (and frictions) into the transferable-utility trading network model of Hatfield et al. (2013).

\(^6\)We rule out intentional-flooding and dynamite-based solutions by assumption.

It costs B 110 to construct a high-end property and 50 to construct a low-end property. It costs each real estate agent 0 to intermediate between B and C. We suppose that C values the high-end property more than the low-end property but experiences income effects—when prices are high, C prefers to buy the low-end property. Intuitively, the high-end property might require additional fees, such as maintenance costs and property taxes, which could cause C to prefer the low-end property when prices are high, but not when prices are low. Moreover, C cannot costlessly convert a high-end property into a low-end property.\(^6\) Formally, C has utility function

\[
\begin{align*}
\begin{array}{c}
\text{C} \\
\text{high-end property}
\end{array} \\
2t + 600 \\
\begin{array}{c}
\text{C} \\
\text{low-end property}
\end{array} \\
t + 400
\end{align*}
\]

As there is only one real estate agent of each type (or, more generally, as we do not assume that there is free entry in the markets for real estate agents), it may be possible for the real estate agents to extract rents. Thus, a competitive equilibrium must specify prices that the consumer faces for each type of property separately from the prices that real estate agents face.

There are two possible equilibrium allocations: either B can construct a high-end property and sell it to C via HighEnd, or B can construct a low-end property and sell it to C via LowEnd.
For example, two competitive equilibria are:

\begin{align*}
\text{HighEnd} & \xrightarrow{115} B \xrightarrow{50} \text{LowEnd} \quad \text{and} \quad \text{HighEnd} & \xrightarrow{152} B \xrightarrow{95} \text{LowEnd}. \\
\text{HighEnd} & \xrightarrow{120} C \xrightarrow{50} \quad \text{and} \quad \text{HighEnd} & \xrightarrow{152} C \xrightarrow{100} \\
\end{align*}

(Here, we denote competitive equilibria by writing a price for each interaction in (6) and boxing the prices of interactions that occur.) In Equilibrium (A), HighEnd extracts a rent of 5 from intermediating between B and C; in Equilibrium (B), LowEnd extracts a rent of 5 from intermediating. However, HighEnd does not trade in Equilibrium (B), while LowEnd does not trade in Equilibrium (A). Thus, there are agents who do not trade in one equilibrium but receive utility strictly greater than their autarky payoffs in the other equilibrium—that is, the lone wolf result fails.

In our example, the builder is able to extract more surplus when the low-end property is traded because the consumer is willing to pay more for the low-end property due to the failure of free disposal. Similarly, the consumer is able to extract more surplus from the builder when the high-end property is traded due to having higher marginal utility of wealth after buying the high-end property. Therefore, the high-end real estate agent is able to improve the consumer’s utility, while the low-end real estate agent improves the builder’s
utility, making it possible for either real estate agent to extract rents.\footnote{Formally, the extremal equilibria are}

In contrast, our Lone Wolf Theorem shows that if utility is transferable, then at most one of the real estate agents can extract rents. Intuitively, if utility is transferable, then the builder and consumer agree on which possible trade is better.\footnote{It might be the case that both trades generate the same social surplus—but when utility is transferable, the builder and the consumer will agree on this fact.} In the presence of income effects, on the other hand, agents can contribute to the social surplus in one allocation but not in another. For example, $\text{HighEnd}$ must contribute to the economy when the high-end property is traded, as she is able to extract rents in Equilibrium (A). However, because $\text{HighEnd}$ does not participate in trade in Equilibrium (B), she cannot possibly contribute to the economy when the low-end property is traded.

5 Discussion and Conclusion

We developed a class of Lone Wolf Theorems that provide equilibrium-independent predictions of competitive equilibrium analysis in contexts with indivisibilities. Our results show that when utility is perfectly transferable, any agent who does not participate in trade in one competitive equilibrium must receive her autarky payoff in every competitive equilibrium; moreover, this result holds approximately under approximate solution concepts and in settings in which utility is only approximately transferable. Partial transferability of utility is
essential for our results—the lone wolf conclusion fails when utility is not transferable (e.g., in the presence of income effects).

5.1 Relationship to the Lone Wolf Theorems and Rural Hospitals

Theorems of Matching Theory

Our lone wolf results extend a classical matching-theoretic insight of McVitie and Wilson (1970) to exchange economies. Other analogues of the Lone Wolf Theorem have been developed in matching, but those results have different hypotheses and conclusions from ours.

The most-general matching-theoretic generalizations—developed by Hatfield and Kominers (2012) and Fleiner, Jankó, Tamura, and Teytelboym (2015)—state that for every agent in a trading network (without transfers), the difference between the numbers of the goods bought and sold is invariant across stable outcomes. Our results extend the Lone Wolf Theorem to exchange economies (with transfers) by assessing when agents participate in (profitable) trade instead of analyzing the amounts that agents trade. Furthermore, the matching-theoretic results rely on two regularity conditions—some form of (gross) substitutability (Kelso and Crawford, 1982), and a regularity condition called the “law of aggregate demand” which we do not require. We instead require that utility is at least approximately transferable.

Recently, Schlegel (2016) has proven a lone wolf result for many-to-one matching with continuous transfers. While Schlegel (2016) allowed workers (i.e., agents on the unit-demand side) to experience income effects, he required that firms (i.e., agents on the multi-unit-demand side) have utility functions that are not only quasilinear but also gross substitutable.

---

9These results are typically called Rural Hospitals Theorems because they imply that that clearinghouses cannot improve the recruitment total of rural hospitals—which are often less attractive to doctors—by switching to different (stable) market-clearing mechanisms. The first Rural Hospitals Theorems were proven by Roth (1984, 1986) for many-to-one matching with responsive preferences. Rural Hospitals Theorems were also developed for many-to-many matching (Alkan (2002); Klijn and Yazıcı (2014)), many-to-one matching with contracts (Hatfield and Milgrom (2005); Hatfield and Kojima (2010)), and many-to-many matching with contracts (Hatfield and Kominers (2017)).

10The law of aggregate demand follows from substitutability in settings with quasilinear utility functions (Hatfield and Milgrom (2005); Hatfield et al. (2019)).
Thus, the Schlegel (2016) lone wolf result is logically independent of ours.

5.2 Application to Strategy-Proofness

It has been well known since the work of Dubins and Freedman (1981) and Roth (1982) that the Gale–Shapley (1962) deferred acceptance mechanism is dominant-strategy incentive compatible for all unit-demand agents on the “proposing” side of the market; this one-sided strategy-proofness result has been important in practice (see, e.g., Roth and Peranson (1999); Abdulkadiroğlu, Pathak, and Roth (2005); Abdulkadiroğlu, Pathak, Roth, and Sönmez (2005); Pathak and Sönmez (2008)) and has been extended to settings with discrete transfers (or other discrete contracts; see Hatfield and Milgrom (2005); Hatfield and Kojima (2010); Hatfield and Kominers (2012, 2019)).

One-sided strategy-proofness results have heretofore been difficult to derive in matching settings with continuous transfers—in part because the now-standard proof of one-sided strategy-proofness (due to Hatfield and Milgrom (2005)) relies on the matching-theoretic Lone Wolf Theorem, and prior to our work there was no lone wolf result for settings with transfers.\textsuperscript{11} In a companion paper (Jagadeesan et al. (2018)), we use the results developed here to give a direct proof of one-sided strategy-proofness for worker–firm matching with continuously transferable utility.\textsuperscript{12}

5.3 Divisible Goods

Finally, we note that while we assumed that the domains of agents’ valuations consist of integer quantity vectors, identical arguments apply if instead agents maximize over real quantity vectors. Thus, although our model formally requires that goods be indivisible, our

\textsuperscript{11} Indeed, it appears that until Hatfield, Kojima, and Kominers (2019) found an indirect argument by way of a version of Holmström’s (1979) lemma, one-sided strategy-proofness of deferred acceptance in the presence of continuously transferable utility was known only for one-to-one markets (Demange (1982); Leonard (1983); Demange and Gale (1985); Demange (1987)).

results apply in settings with divisible goods as well.
A Proof of Lemma 1

As \([\hat{q}; \hat{p}]\) is an \(\varepsilon\)-equilibrium, we have that

\[
\varepsilon \geq \sum_{i \in I} \left( \max_{q^i \in \mathbb{Z}} \left\{ u^i(\hat{q}^i, (e^i - q^i) \cdot \hat{p}) \right\} - u^i(\hat{q}^i, (e^i - \hat{q}^i) \cdot \hat{p}) \right)
\]

\[
\geq \sum_{i \in I} \left( u^i(q^i, (e^i - q^i) \cdot \hat{p}) - u^i(\hat{q}^i, (e^i - \hat{q}^i) \cdot \hat{p}) \right)
\]

\[
= \sum_{i \in I} \left( v^i(q^i) - v^i(\hat{q}^i) \right) + \left( \sum_{i \in I} q^i - \sum_{i \in I} \hat{q}^i \right) \cdot \hat{p}
\]

\[
= \sum_{i \in I} \left( v^i(q^i) - v^i(\hat{q}^i) \right),
\]

where the last equality holds as \(\sum_{i \in I} q^i = \sum_{i \in I} e^i = \sum_{i \in I} \hat{q}^i\). Hence, we have that

\[
\sum_{i \in I} v^i(\hat{q}^i) + \varepsilon \geq \sum_{i \in I} v^i(q^i).
\]

It follows that

\[
\sum_{i \in I} u^i(q^i, (e^i - q^i) \cdot p) = \sum_{i \in I} v^i(q^i) + \sum_{i \in I} \left( e^i - q^i \right) \cdot p
\]

\[
= \sum_{i \in I} v^i(q^i)
\]

\[
\leq \sum_{i \in I} v^i(\hat{q}^i) + \varepsilon
\]

\[
= \sum_{i \in I} v^i(\hat{q}^i) + \sum_{i \in I} \left( e^i - \hat{q}^i \right) \cdot p + \varepsilon
\]

\[
= \sum_{i \in I} u^i(\hat{q}^i, (e^i - \hat{q}^i) \cdot p) + \varepsilon. \tag{7}
\]

Grouping (7) by agents, we have

\[
\sum_{i \in I} \left( u^i(q^i, (e^i - q^i) \cdot p) - u^i(\hat{q}^i, (e^i - \hat{q}^i) \cdot p) \right) \leq \varepsilon. \tag{8}
\]

\[^{13}\text{Intuitively, this argument shows that the allocation } (\hat{q}^i)_{i \in I} \text{ must be within } \varepsilon \text{ utils of maximizing the sum of agents' values over all allocations—i.e., that the allocation is approximately efficient.}\]
As \([q; p]\) is a \(\delta\)-equilibrium, we have

\[
\sum_{i \in I} \left( \max_{q_i^i \in Z_i} \left\{ u^i(q_i^i, (e_i - q_i^i) \cdot p) \right\} - u^i(q_i^i, (e_i - q_i^i) \cdot p) \right) \leq \delta. \tag{9}
\]

Summing (8) and (9), we have

\[
\sum_{i \in I} \left( \max_{q_i^i \in Z_i} \left\{ u^i(q_i^i, (e_i - q_i^i) \cdot p) \right\} - u^i(q_i^i, (e_i - q_i^i) \cdot p) \right) \leq \delta + \varepsilon.
\]

Hence, \([\hat{q}; p]\) is a \((\delta + \varepsilon)\)-equilibrium, as claimed.
References


