# Decentralized Matching with Aligned Preferences*

Andrew Ferdowsian[†] Muriel Niederle [‡] Leeat Yariv[§]

*This version*: February 14, 2022, New Version

**Abstract**

We study a simple model of a decentralized market game in which firms make directed offers to workers. We focus on markets in which agents have aligned preferences. When agents have complete information or when there are no frictions in the economy, there exists an equilibrium that yields the stable matching. In the presence of market frictions and preference uncertainty, harsher assumptions on the economy's richness are required for decentralized markets to generate stable outcomes in equilibrium.

**Keywords:** Decentralized Matching, Stability, Market Design.

**JEL:** C72, C73, C78, D47, D83.

*We thank Attila Ambrus, Federico Echenique, Guillaume Haeringer, Fuhito Kojima, and Al Roth for many useful conversations and suggestions. Niederle gratefully acknowledges support from the Sloan Foundation. Yariv gratefully acknowledges support from NSF grants SES-0963583 and SES-1629613.

[†]Department of Economics, Princeton University.

[‡]Department of Economics, Stanford University and NBER.

[§]Department of Economics, Princeton University, CEPR, and NBER.

# 1 Introduction

## 1.1 Overview

The theoretical literature on two-sided matching markets has focused predominantly on the analysis of outcomes generated through centralized clearinghouses. Examples of such markets include the medical residency match, school allocations, etc. Nonetheless, many markets are not fully centralized and have participants interacting in a less structured fashion over time: college admissions in the U.S., the market for law clerks, junior economists, and more. Furthermore, almost all centralized markets are preceded by decentralized opportunities for participants to match.[1] Understanding the outcomes generated by decentralized markets is therefore important for the design of decentralized and centralized institutions alike. This paper offers a step in that direction.

Empirically, persistent centralized clearinghouses are associated with stable outcomes, see Roth (1991).[2] A common folk argument contends that if an unstable outcome is implemented, decentralized interaction, either preceding or following the operation of the clearinghouse, would ultimately yield stability. Such decentralized interactions may come at an efficiency cost when frictions are present. For example, in labor markets, employers and employees already matched, who potentially have preferable matches elsewhere, would need to seek those out and take costly steps to undo their current arrangements. In this paper, we inspect theoretically this folk wisdom. We provide conditions under which (non-cooperative) decentralized market interactions yield stability. In particular, we illustrate how both time and information frictions not only affect outcomes' efficiency, but also introduce critical obstacles to stability. Methodologically, our characterization offers a non-cooperative foundation for the cooperative stability notion.

We suggest a simple model of a decentralized *market game* in which firms make di-

[1] See Roth (1984) and Roth (2008). For some evidence on outcome differences between centralized and decentralized markets, see Fréchette, Roth, and Ünver (2007) and Niederle and Roth (2003). These papers also show that the consequences of a decentralized matching process prior to a centralized match can be large, as documented by the collapse of the market for gastroenterology fellows.

[2] A *stable matching* is a pairing of workers and firms (where some workers and some firms may remain unmatched), in which no firm (worker) who is matched to a worker (firm), prefers to be alone, and no firm and worker pair prefer to jointly deviate by matching to one another.

rected offers to workers. A market game is identified by three components: the preference distribution of agents (workers and firms), the information agents have about their own and others' realized preferences, and the extent of time frictions in the economy.

For simplicity, we focus on markets in which firms can employ up to one worker, who can work for at most one firm, and assume that a match with any agent is preferred to no match at all. We consider environments in which there is a unique stable matching. We do so for two reasons. First, uniqueness allows us to sidestep coordination problems, thereby eliminating one straightforward hurdle to stability. Second, there is work suggesting a small number of stable matchings in some settings, see Roth and Peranson (1999). In fact, recent theoretical work suggests that, under some conditions, large markets have a small number of stable matchings (see Immorlica and Mahdian 2005, Kojima and Pathak 2009, and Ashlagi, Kanoria, and Leshno 2017). As many of the leading applications of matching theory—from labor markets to school allocations—involve tens or hundreds of thousands of participants, these results are in line with our uniqueness assumption.

We concentrate our analysis on markets with *aligned preferences*. Aligned preferences require that the preferences of firms and workers can be represented by a joint "ordinal potential," analogous to ordinal potentials of normal-form games (Monderer and Shapley, 1996). An ordinal potential specifies a value for each firm and worker *pair* and captures the full preference profile of market participants. Aligned preferences guarantee uniqueness of the stable matching. Many prominent cases studied in the literature entail aligned preferences. For instance, alignment is guaranteed whenever firms and workers split revenue in fixed proportions, or when all participants on at least one side of the market share the same preference ranking of the other side's participants.

In our decentralized market game, in every period, each firm can make up to one offer if she does not already have an offer held by some worker.[3] Workers can accept, reject, or hold on to an offer, where holding on to an offer ensures its availability in the future. To study the impacts of time frictions, we assume firms and workers receive discounted match utilities when they are matched.

We study two settings that differ in the information agents have about their own and

[3] Throughout, we refer to firms as female and workers as male.

others' preferences. Under *complete information*, all market participants are fully informed of the realized match utilities. While the extant literature has mostly focused on complete information environments (for a few exceptions, see the literature review below), allowing for incomplete information is particularly important from an applied perspective. Many matching markets are large and entail limited communication between participants. For instance, the job market for economists includes several thousands of applicants and new positions every year, the National Resident Matching Program (NRMP) involves around 70,000 participants annually, etc. Participants cannot possibly communicate their complete preference rankings to all others. Our *incomplete information* environment aims to capture such scenarios: each agent is fully informed only of her own match utilities and holds (correct) prior belief distributions pertaining to others' preferences.

We analyze equilibria in weakly undominated strategies. Under complete information, all agents can compute the stable matching, and it is an equilibrium outcome of the decentralized market game. Still, there may be equilibria that yield unstable matchings, highlighting a first contrast with findings from the case in which agents participate in a centralized market, where all equilibria in weakly undominated strategies yield the stable matching. At the crux of this multiplicity is the observation that in dynamic decentralized markets, agents can condition their actions on past activity. Nonetheless, a simple refinement—namely, iterated elimination of weakly dominated strategies—restores uniqueness of the stable matching as an equilibrium outcome.[4]

With incomplete information, in a frictionless economy, stable matchings can still be implemented as an equilibrium outcome. Underlying this result is the idea that firms and workers can replicate, in essence, the firm-proposing Gale-Shapley deferred acceptance algorithm (Gale and Shapley, 1962) as part of an equilibrium profile. That is, firms make offers to workers in order of their preferences, and workers accept offers when they are made by their most preferred firms. Aligned preferences ensure there is always a firm and a worker that are each other's favorite partner. This guarantees that at least one firm and worker are matched in every period. In particular, such a strategy profile leads to a stable

[4]This is consistent with laboratory observations of decentralized markets with complete information, in which stability is frequently achieved Echenique, Robinson-Cortes, and Yariv (2012) and Niederle and Roth (2009).

market outcome in finite time.

Nevertheless, even with small time frictions, agents may have incentives to deviate from these strategies. Intuitively, if all firms and workers emulate the firm-proposing deferred acceptance algorithm, the timing of events—offers, responses to offers, and exits—feeds into the updating process. There are then two types of manipulations: ones intended to speed up the matching process, and ones intended to affect participants' learning regarding their expected stable matches.

To glean intuition for manipulations driven by a desire to speed up matches, consider an economy with two workers, Adrian and Bailey, and two firms, $A$ and $B$. Suppose that (i) both workers prefer firm $A$ to firm $B$ and (ii) both firms prefer Bailey to Adrian. Under deferred acceptance strategies, both firms approach Bailey, their favorite, first. Adrian then receives an offer from firm $B$ only in the second period. If firm $B$ proceeds out of order and makes an offer to Adrian in the first period, Adrian would accept that offer immediately. In this case, firm $B$ successfully *speeds up* her match. In a richer example, such speeding motives can yield unstable matchings.

To understand what underlies manipulations designed to alter matches, consider a different economy. Suppose that, using deferred acceptance strategies, some agent, say Cooper, matches with some firm $C$ only when firm $C$ makes an offer to Cooper in the first period. This provides firm $C$ incentives to make an early offer to Cooper in markets in which Cooper is preferable to her stable match partner. Making an early offer to Cooper allows firm $C$ to manipulate Cooper's beliefs about the realized market and secure a superior match.

In order to implement stable outcomes through equilibrium play, we need to limit the learning that occurs from the timing of events per se. Certainly, when information is complete, there are no learning opportunities and, indeed, the stable outcome can be implemented in equilibrium. The second step of our analysis illustrates that when the economy is sufficiently rich, namely when *any* aligned preference profile is possible with some positive probability, learning is restricted enough for implementation when agents are sufficiently patient. In rich economies, there are many ways to rationalize various timing of events through the supported preference profiles. Consequently, timing plays a more limited role in learning.[5]

Since many applications pertaining to our model involve numerous participants, one may worry that the equilibria we identify demand a long time for the market to stabilize. Using simulations of randomly generated preferences, we show that the number of periods required for markets of reasonable size to culminate in a stable outcome is limited.

### 1.2 Related Literature

Several recent papers study decentralized interactions in matching markets empirically, and indicate the relevance of incomplete information for outcomes. Grenet, He, and Kübler (2021) consider a decentralized phase in Germany's university admission system and find evidence for students' costly learning about universities. Echenique et al. (2022) show the importance of the decentralized interview process preceding the centralized medical match in the US. Using experiments, Pais, Pintér, and Veszteg (2012), Echenique, Robinson-Cortes, and Yariv (2012), and Agranov et al. (2021) consider selection of and convergence to stable matchings in lab decentralized markets (see also references therein). While lab complete-information markets often culminate in stable matchings, that is not the case with incomplete information.

A few theoretical investigations inspect market outcomes as consequences of a dynamic interaction. Haeringer and Wooders (2011) and Pais (2008) consider the case of complete information, and restrict firms' strategies by preventing offers to workers who had rejected them previously. Haeringer and Wooders (2011) study a game similar to ours in which firms can only make exploding offers (that have to be accepted or rejected right away). Pais (2008) investigates a setting with one firm chosen randomly each period to make an offer. Pais characterizes the set of (ordinal) subgame perfect equilibria and shows that outcome multiplicity may arise even when the underlying market has a unique stable matching. Suh and Wen (2008) consider a particular sequential protocol of offers in which each participant makes only one decision—accept a previous offer, stay single, or make an offer to someone who follows them. With complete information, they show conditions under which subgame perfect equilibrium outcomes are stable. We add to this line of

[5] In the Appendix, we provide more general sufficient conditions for implementation of stable matchings.

work by considering markets in which both firms and workers interact strategically. We also consider both complete- and incomplete-information settings.[6]

Most of the theoretical matching literature effectively assumes complete information of preferences. Liu et al. (2014) and Liu (2020) suggest a cooperative notion of stability with incomplete information allowing for transfers, while Bikhchandani (2017) offers a notion absent transfers. Roth (1989) and Fernandez, Rudov, and Yariv (2022) illustrate the impacts of incomplete information in a centralized matching clearinghouse.[7] We contribute to this literature by suggesting non-cooperative equilibration dynamics. We identify environments in which, even with incomplete information, the complete-information stable outcomes emerge as equilibrium outcomes.

The recent and growing literature on dynamic matching also considers dynamic interactions, see e.g., Akbarpour, Li, and Gharan (2020), Baccara, Lee, and Yariv (2020), and the survey by Baccara and Yariv (2022). In that literature, market participants arrive over time—as is the case for organ patients and donors, children relinquished for adoption and parents seeking to adopt, etc. The focus is often on optimal clearinghouses for such settings, rather than outcomes of decentralized interactions we consider.

The search and matching literature—Burdett and Coles (1997), Eeckhout (1999), Shimer and Smith (2000), and the survey by Chade, Eeckhout, and Smith (2017)—studies dynamic interactions in an often stationary market in which agents are commonly evaluated. As in

[6] Roth and Vande Vate (1990) offer non-strategic dynamics of blocking-pair formation that yield stability. Dai and Jordan (2021) take an algorithmic approach for learning procedures that approximate payoff-maximizing strategies in decentralized markets. Arnosti, Johari, and Kanoria (2014) study decentralized matching markets in which agents arrive and depart asynchronously. They study the effects of observing who is available. The role of commitment in dynamic games such as ours with complete information is highlighted in Diamantoudi, Miyagawa, and Xue (2015). Similarly, Blum, Roth, and Rothblum (1997) study dynamics when the firms' but not the workers' commitments are binding. There is also some work analyzing endogenous salaries in decentralized markets with complete information and limited dynamics, see Konishi and Sapozhnikov (2008). The link between dynamic interaction and stability has been suggested in the context of implementation as well. With complete information, Alcalde and Romero-Medina (2000) study a game of two stages. First, firms make offers. Then, workers reply. They demonstrate that this game implements the stable matchings Alcalde, Pérez-Castrillo, and Romero-Medina (1998).

[7] Ehlers and Massó (2007) also consider incomplete information in stable mechanisms. They link singleton cores and truthful preference revelation as an ordinal Bayesian Nash equilibrium.

our setting, time frictions present a hurdle to stability: they impose a search cost that leads agents to "compromise" on match partners. However, in these models, as time frictions vanish, stability is approximated. We consider a richer class of preferences and show that, with incomplete information, even vanishingly small time frictions can present a severe impediment to stability unless learning opportunities are limited.

Gutin, Neary, and Yeo (2021) characterize preference restrictions guaranteeing a unique stable matching. Our alignment assumption is somewhat more restrictive, but reminiscent of some identified sufficient conditions for uniqueness (Clark 2006 and Eeckhout 2000). Some of its special cases often appear in empirical work. For example, a similar assumption is made in Agarwal (2015) when analyzing the medical residency match and in Dur et al. (2018) when examining a school-choice setting in Taiwan.

# 2 The Model

## 2.1 The Economy

A *market* is a triplet $\mathcal{M} = (\mathcal{F}, \mathcal{W}, U)$, where $\mathcal{F} = \{1, ..., F\}$ and $\mathcal{W} = \{1, ..., W\}$ are disjoint sets of firms and workers, respectively, and $U = \left\{\left\{u_{ij}^f\right\}, \left\{u_{ij}^w\right\}\right\}$ are agents' match utilities.[8] Each firm $i \in \mathcal{F}$ has match utility $u_{ij}^f$ from matching to worker $j \in \mathcal{W} \cup \varnothing$, where matching to $\varnothing$ is interpreted as remaining unmatched. Similarly, for each worker $j \in \mathcal{W}$, $u_{ij}^w$ is the match utility from matching to firm $i \in \mathcal{F} \cup \varnothing$. We denote by $U^f = \left(u_{ij}^f\right)_{i \in \mathcal{F}, j \in \mathcal{W}}$ and $U^w = \left(u_{ij}^w\right)_{i \in \mathcal{F}, j \in \mathcal{W}}$ the matrices corresponding to utilities from firm-worker pairs for both sides of the market.

For simplicity, we assume that all match utilities are strictly positive, that firms and workers have strict preferences, and that all agents prefer to be matched over remaining unmatched.[9] That is, for any firm $i$ and workers $j, j' \in \mathcal{W} \cup \varnothing$,

$$u_{ij}^f \neq u_{ij'}^f > u_{i\varnothing}^f > 0.$$

Similarly, for any worker $j$ and firms $i, i' \in \mathcal{F} \cup \varnothing$,

[8] Cardinal utilities are necessary for trading off matches at different points in time and examining the impacts of discounting.

[9] Allowing for some agents to be unacceptable would hurt presentation without altering our main messages.

$$u^w_{ij} \neq u^w_{i'j} > u^w_{\varnothing j} > 0.$$

In particular, all outside options generate strictly positive utilities.[10]

For fixed sets $\mathcal{F}$ and $\mathcal{W}$ of firms and workers, an *economy* is a finite collection of markets $\{(\mathcal{F}, \mathcal{W}, U)\}_{U \in \mathcal{U}}$ together with a distribution $G$ over possible match utilities $U \in \mathcal{U}$. We assume $G$ has finite support $\mathcal{U}$: each market in the economy has positive probability and cannot be ignored in agents' considerations.

A *matching* is a function $\mu : \mathcal{F} \cup \mathcal{W} \to \mathcal{F} \cup \mathcal{W} \cup \varnothing$ such that for all $i \in \mathcal{F}$, $\mu(i) \in \mathcal{W} \cup \varnothing$, and for all $j \in \mathcal{W}$, $\mu(j) \in \mathcal{F} \cup \varnothing$. Furthermore, if $(i, j) \in \mathcal{F} \times \mathcal{W}$, then $\mu(i) = j$ if and only if $\mu(j) = i$. If $\mu(k) \neq \varnothing$ for $k \in \mathcal{F} \cup \mathcal{W}$, we say that $k$ is matched under $\mu$. A *blocking pair* for a matching $\mu$ is a pair $(i, j) \in \mathcal{F} \times \mathcal{W}$ such that $u^f_{ij} > u^f_{i\mu(i)}$ and $u^w_{ij} > u^w_{\mu(j)j}$. A matching is *stable* if there exist no blocking pairs. Since all agents prefer to be matched, individuals never block a matching.

As noted in our literature review, in a static setting, one could consider an adjusted notion of stability that allows for incomplete information; see, e.g., Liu et al. (2014) and Bikhchandani (2017). Our goal here is to understand when learning in matching markets can generate the complete-information stable outcomes. When it does, outcomes are expected to be robust to any further information or interactions after the dust has settled in the market. Throughout the paper, we often slightly abuse terminology and refer to the complete-information stable matchings in a market as simply stable matchings.

Gale and Shapley (1962) showed that any market has a stable matching, and provided an algorithm that identifies one. In the *firm-proposing deferred acceptance (DA) algorithm*, in step 1, each firm makes an offer to its most preferred worker. Workers collect offers, hold the offer from their most preferred firm, and reject all other offers. In a general step $k$, each firm whose offer was rejected in the last step makes an offer to the most preferred worker who has not rejected her yet. Workers once more collect offers including, possibly, an offer held from a previous step, keep their most preferred offer, and reject all other offers. The algorithm ends when there are no more offers that are rejected; that is, when

[10] With an outside option that generates a utility of 0, agents expecting to exit do not care about the timing of their exit, which would generate equilibrium multiplicity in a trivial manner.

each firm either has her offer held by a worker, or has been rejected by all workers. Once the algorithm ends, held offers convert to matches.

While the firm-proposing DA algorithm generates one stable matching, the one preferred by all firms, there may be others. Certainly, in markets with multiple stable matchings, learning to coordinate on a particular (complete-information) stable matching for any market realization would entail many difficulties. For one, agents need to learn which stable matching is being played, in addition to various relevant features of the market. Furthermore, incentive compatibility concerns akin to those emerging in centralized settings with multiple stable matchings may surface in decentralized markets as well (see Roth and Sotomayor 1992, and Section 2.4 below). In order to circumvent these hurdles and isolate the impacts of incomplete information and frictions on learning in matching markets, we assume that any market $\mathcal{M} = (\mathcal{F}, \mathcal{W}, U)$ in the support of $G$ has a unique stable matching denoted by $\mu_M$. This assumption is in line with prior work. For instance, Roth and Peranson (1999)'s analysis of the medical residents' match data finds small cores. As already mentioned, recent theoretical work (see Immorlica and Mahdian 2005, Kojima and Pathak 2009, and Ashlagi, Kanoria, and Leshno 2017) supports the idea that, in applications that involve a large participant volume, there is a small number of stable matchings.

## 2.2 Aligned Preferences

We focus on a class of preferences we term *aligned preferences* that exhibit a unique stable matching and several other desirable features.

**Definition** (Aligned Preferences)**.** Firms and workers have *aligned preferences* if there exists an *ordinal potential* $\Phi = \left(\Phi_{ij}\right)_{i\in\mathcal{F}, j\in\mathcal{W}}$, $\Phi_{ij} \in \mathbb{R}$ such that for any workers $j, j' \in \mathcal{W}$ and firms $i, i' \in \mathcal{F}$:

$$\text{If } u^w_{ij} > u^w_{i'j}, \text{ then } \Phi_{ij} > \Phi_{i'j}, \text{ and if } u^f_{ij} > u^f_{ij'}, \text{ then } \Phi_{ij} > \Phi_{ij'}.$$

Intuitively, preference alignment imposes a link between firms' and workers' (ordinal) preferences through the ordinal potential. The preference ranking of both sides of the market can be captured using one common matrix $\Phi$. Let $\hat{U}^f = \hat{U}^w = \left(\Phi_{ij}\right)_{i\in\mathcal{F}, j\in\mathcal{W}}$. Then, $\hat{U}^f$ and $\hat{U}^w$ capture the same ordinal preferences over partners as $U^f, U^w$.[11] Many applied

[11]The notion of ordinal potential is analogous to that of a potential in two-player games in which agents'

papers implicitly assume that preferences are aligned (e.g., Sørensen 2007, Agarwal 2015, and Dur et al. 2018).

Several special cases of preference alignment are prominent in the literature. For instance, suppose firms and workers have a joint production output they share in fixed proportions when matched. That is, for $\alpha > 0$ and all $(i, j) \in \mathcal{F} \times \mathcal{W} : u^w_{ij} = \alpha u^f_{ij} > 0$. Here, $\Phi = \left(\Phi_{ij}\right)_{i \in \mathcal{F}, j \in \mathcal{W}}$ defined as $\Phi_{ij} = u^f_{ij}$ for all $i, j$ serves as an ordinal potential. Alternatively, suppose participants on one side of the market rank participants on the other side of the market identically—firms may all value potential employees using their college GPA, or students applying to colleges may all use the *U.S. News and World Report* rankings to form their preferences. Such preferences are aligned as well.[12]

As we now describe, preference alignment implies uniqueness of the stable matching, in addition to several properties useful for our analysis.

Define a market $(\tilde{\mathcal{F}}, \tilde{\mathcal{W}}, \tilde{U})$ as a *submarket* of $(\mathcal{F}, \mathcal{W}, U)$ if $\tilde{\mathcal{F}} \subseteq \mathcal{F}$, $\tilde{\mathcal{W}} \subseteq \mathcal{W}$, and $\forall i, j \in (\tilde{\mathcal{F}} \cup \varnothing) \times (\tilde{\mathcal{W}} \cup \varnothing) \setminus \{\varnothing, \varnothing\}$, $\tilde{u}^w_{ij} = u^w_{ij}$ and $\tilde{u}^f_{ij} = u^f_{ij}$. A direct implication of preference alignment is that, for any submarket $(\tilde{\mathcal{F}}, \tilde{\mathcal{W}}, \tilde{U})$, there is a firm $i$ and a worker $j$ that form a *top-top* match, i.e. worker $j$ is firm $i$'s most preferred worker within $\tilde{\mathcal{W}}$ and firm $i$ is worker $j$'s most preferred firm within $\tilde{\mathcal{F}}$. Indeed, when preferences are aligned, the original market, as well as any submarket, has an ordinal potential. Suppose $\tilde{\Phi}$ is an ordinal potential of the submarket $(\tilde{\mathcal{F}}, \tilde{\mathcal{W}}, \tilde{U})$. Consider a pair $(i, j) \in \arg\max_{(i', j') \in \tilde{\mathcal{F}} \times \tilde{\mathcal{W}}} \tilde{\Phi}_{i'j'}$. It follows that $(i, j)$ is a top-top match in $(\tilde{\mathcal{F}}, \tilde{\mathcal{W}}, \tilde{U})$. When every submarket has a top-top match, we say that preferences satisfy the **top-top match property**. Preference alignment implies the top-top match property.[13]

The top-top match property guarantees uniqueness of the stable matching. Indeed, find the firm-worker pairs that constitute top-top matches, pairs $(i, j)$ at which the ordinal

match utilities replace the payoff matrix $\left(U^w, U^f\right) = \left((u^w_{ij}, u^f_{ij})\right)_{i \in \mathcal{F}, j \in \mathcal{W}}$ Monderer and Shapley (1996).

[12]Suppose, without loss of generality, that firms share a single preference ranking over workers. Order the workers according to this ranking, with worker 1 being the least preferred worker and worker $n$ being the most preferred worker. Normalize each worker's match utilities to $\tilde{u}^w_{ij}$, such that for all $i, k \in \mathcal{F}, j \in \mathcal{W}$: $0 < \tilde{u}^w_{ij} < 1$, and $\tilde{u}^w_{ij} < \tilde{u}^w_{kj} \Leftrightarrow u^w_{ij} < u^w_{kj}$. Then $\Phi = \left(\Phi_{ij}\right)_{i \in \mathcal{F}, j \in \mathcal{W}}$ with $\Phi_{ij} = j + \tilde{u}^w_{ij}$ for all $i, j$ is an ordinal potential.

[13]The top-top match property is referred to as $\alpha$-reducibility in Clark (2006). The top-top match property does not imply alignment on its own.

potential $\Phi = \left(\Phi_{ij}\right)$ is maximized. A firm and a worker in such a pair are each other's favorites in the market and therefore must be matched in any stable matching. The remaining firms and workers form a submarket with aligned preferences. We can continue recursively to achieve the unique stable matching.

Suppose $\mu_M$ is the unique stable matching of market $M = (\mathcal{F}, \mathcal{W}, U)$ with aligned preferences. The **stable blocking pair property** holds. That is, for any unstable matching $\mu$, there exists a *stable blocking pair*: a blocking pair $(i, j)$ for which $\mu_M(i) = j$. Indeed, suppose $\mu \neq \mu_M$. Consider the recursive process described above to illustrate uniqueness. At some stage, a discrepancy must arise between $\mu_M$ and $\mu$. At that stage, a match that occurs under $\mu_M$ does not get formed. The corresponding worker and firm form a stable blocking pair.

Finally, preference alignment implies that when firms make offers in the order of their preferences, a rejected offer of a worker cannot trigger a chain of offers and rejections that results in an offer from a more desirable firm. That is, if a worker $j$ rejects an offer from firm $i$, the resulting chain of offers can only result in offers to worker $j$ that he prefers less than the offer from firm $i$. Formally, there is no sequence $i_1, ..., i_n \in \mathcal{F}$ and $j_1, ..., j_n \in \mathcal{W}$ such that, by rejecting firm $i_2$, worker $j_1$ can trigger an offer from a preferred firm $i_1$ :

$$u^w_{i_1 j_1} > u^w_{i_2 j_1}, \;\; u^f_{i_2 j_1} > u^f_{i_2 j_2}, \;\; u^w_{i_2 j_2} > u^w_{i_3 j_2}, \;\; u^f_{i_3 j_2} > u^f_{i_3 j_3}, ..., \;\; u^w_{i_n j_n} > u^w_{i_1 j_n}, \;\; u^f_{i_1 j_n} > u^f_{i_1 j_1}$$

Such a chain would be equivalent to having a cycle in the payoff matrix $\left(U^w, U^f\right)$. We thus say that preferences satisfy the **no-cycle property**. The characterization of potential games by Voorneveld and Norde (1997) ensures that preference alignment is *equivalent* to the no-cycle property.

Proposition 1 summarizes the properties of markets with preference alignment.

**Proposition 1** (Alignment – Properties)**.**

1. *If preferences are aligned, the stable matching is unique. Furthermore, the top-top match and stable blocking pair properties hold.*

2. *Preferences are aligned if and only if the no-cycle property holds.*

### 2.3 A Decentralized Market

For a given economy $\{(\mathcal{F}, \mathcal{W}, U)\}_{U \in \mathcal{U}}$ and a distribution $G$ over utility realizations (or markets), we analyze the following *market game*. The economy, together with the distribution $G$, are common knowledge to all agents. At the outset, the market is realized according to the distribution $G$. Each agent is privately informed only of their own realized match utilities. When the support of $G$ is a singleton, there is *complete information*: firms' and workers' matching utilities are common knowledge. This is the case that most of the literature on decentralized matching has tackled. With complete information, both firms and workers can deduce the stable outcome. We also analyze settings in which the support of $G$ is non-trivial. In this *incomplete information* case, participants may not be able to infer the underlying market, nor its stable matching, from their private information.

In our market game, firms make offers over time, indexed by $t \in \{1, 2, ...\}$, and workers react to those offers. Specifically, each period has three stages. In the first stage, eligible firms simultaneously decide whether and to whom to make an offer and whether to exit the market.In the second stage of any period, workers observe which firms exited, and observe the offers they themselves received. Each worker $j$ who has received an offer from firm $i$ can accept, reject, or hold the offer. If an offer is accepted, worker $j$ is matched to firm $i$. Workers can also exit the market. In the third stage, firms observe rejections and deferrals of their own offers. Finally, all participants are informed of the agents who exited the market and the participants who were matched.[14]

Eligible firms are firms that have not yet hired a worker and have no offer held by a worker. In each period $t$, eligible firms can make up to one offer to any worker that has not yet been matched. The game ends once all participants have exited the market, matched or unmatched.

We consider market games with and without time frictions, which take the form of discounting. If a firm $i$ is matched to worker $j$ at time $t$, firm $i$ receives $\delta^t u_{ij}^f$ and worker $j$ receives $\delta^t u_{ij}^w$, where $\delta \in [0, 1]$ is the market discount factor. As long as agents are unmatched, they receive 0 in each period. One interpretation is that once a worker and a

[14]Participants do not observe the entire stream of offers and responses made by others, only their matches and exits. This assumption fits many applications and helps with tractability of the belief-updating process.

firm are matched, they receive their match utility or, equivalently, they receive a constant, perpetual stream of payoffs, the present value of which is their match utility. One can also interpret the discount factor $\delta$ as the probability of market collapse, or the probability that each firm loses its position and receives a payoff of 0 (and, analogously, the probability that each worker leaves the market and receives 0 as well).[15]

We assume all offer benefits are captured by match utilities. In particular, firms do not offer targeted transfers or wages to workers. Hall and Krueger (2012) suggests that a substantial fraction of jobs entails posted wages, plausibly resistant to general equilibrium forces. Our analysis pertains to such settings.[16]

The equilibrium notion we use is that of Bayesian Nash equilibrium, where types correspond to agents' private information. When the support of $G$ is a singleton and information is complete, our analysis pertains to the Nash equilibria of the corresponding game.

We focus on equilibria of decentralized market games in which all agents use weakly undominated strategies. Ruling out weakly dominated strategies imposes several restrictions on equilibrium play:

1. A worker who accepts an offer always accepts his best available offer. In particular, a worker cannot exit and simultaneously reject an offer since offers always generate higher payoffs.

2. When $\delta < 1$, a worker who receives an offer from his most preferred unmatched firm accepts it immediately. Similarly, if only agents on one side of the market are unmatched, they exit immediately since all payoffs, including those from the outside option, are strictly positive.

### 2.4 Centralized Matching Benchmark

Our decentralized market game shares features with the firm-proposing DA algorithm: firms make at most one offer at any time and operate in stages; workers can hold on to of-

[15] An alternative source of frictions would be costly offers. An analogous model in which each offer costs a constant amount would yield qualitatively similar results to those presented here.

[16] With incomplete information, targeted transfers may introduce additional obstacles to stability. Market participants then need to infer not only their correct match partner, but also the appropriate wage.

fers until they are content with one that they accept. There are important differences since market participants' strategies are not restricted. Firms need not go in order of their preference ranking and can make repeat offers; workers can hold multiple offers, and accept or reject any of them. Furthermore, dynamic interactions allow for non-trivial conditioning on and learning from observed histories. Nonetheless, the centralized clearinghouse utilizing the firm-proposing DA is a natural benchmark for our analysis.

To ease analogies between centralized and decentralized markets, we assume that agents in a centralized market report match utilities that are then converted into ordinal preferences. That is, each agent submits a vector of positive match utilities.[17]

Formally, for each type of agent $\alpha \in \{f, w\}$, and each agent $l$, let $P(u_l^\alpha)$ be the strict ordinal preferences associated with $l$'s reported match utilities, in which ties are broken depending on the index of the relevant match partners and in favor of being matched.[18] Specifically, consider firm $i$, then $u_{ij}^f > u_{ik}^f$ implies $jP(u_{i\cdot}^f)k$, for $j,k \neq \varnothing$, $u_{ij}^f = u_{ik}^f$ and $j < k$ implies $jP(u_{i\cdot}^f)k$, and for $j \neq \varnothing$, $u_{ij}^f \geq u_{i\varnothing}^f$ implies $jP(u_{i\cdot}^f)i$.

We define a *deferred acceptance (DA) mechanism* as a mechanism in which all agents report their match utilities simultaneously, after receiving private information. The mechanism then computes the corresponding ordinal preferences as above and outputs the stable matching induced by the firm-proposing DA algorithm. The payoffs of firms and workers are the match utilities corresponding to the implemented matching.

It follows directly from incentive compatibility attributes of the DA algorithm that the DA mechanism allows for a Bayesian Nash equilibrium in weakly undominated strategies in which the resulting matching corresponds to the unique stable matching in each market of the economy (see Roth and Sotomayor 1992). That is,

**Lemma 1** (Centralized Matching)**.**

1. *For any complete-information economy, all Nash equilibria in weakly undominated strategies of the game associated with the DA mechanism yield the unique stable matching.*

2. *For any economy, there exists a Bayesian Nash equilibrium in weakly undominated strate-*

[17]The restriction to positive numbers is made only for presentation simplicity. It suffices that the set of available reports contains as many elements as the maximal number of different match utility vectors.

[18]Although agents are never indifferent in their realized match utilities, they may still report indifferences.

*gies of the game associated with the DA mechanism that yields the unique stable matching in each market.*

Although implementing the stable matching is possible through a centralized clearinghouse whenever there is a unique stable matching (even absent alignment), the stable matching is not necessarily the unique equilibrium outcome in the presence of uncertainty Fernandez, Rudov, and Yariv (2022).

## 3 Decentralized Complete Information Economies

We now turn to our decentralized market game. We start by analyzing economies in which all participants are informed of the details of the market, when the support of $G$ is a singleton. In this case, all agents can compute the stable matching. The stable matching can then be achieved in one period. Intuitively, consider the following strategy profile. Each firm makes an offer to her stable match partner under the unique stable matching $\mu_M$ and exits the market if unmatched under $\mu_M$. Each worker accepts his best available offer in period 1, and exits upon receiving no offers. This profile constitutes an equilibrium in weakly undominated strategies that yields the matching $\mu_M$.

Ruling out weakly dominated strategies is not sufficient to guarantee uniqueness, however. First, there may be multiple equilibria generating $\mu_M$. Indeed, for sufficiently high discount factors, an alternative way of implementing $\mu_M$ through equilibrium involves emulating the DA algorithm. Since this profile may entail several periods of market activity, it can generate different equilibrium payoffs for non-trivial discount factors.

Furthermore, there may be outcomes generated by equilibria in weakly undominated strategies that are unstable, as the following example illustrates.

**Example 1 (Multiplicity with Complete Information)** Assume an economy with four firms $\{F1, F2, F3, F4\}$ and four workers $\{W1, W2, W3, W4\}$, in which $u^f_{ij} = u^w_{ij}$. The following matrix defines the payoffs of all matches:

$$U^f = U^w = \begin{array}{|c|c|c|c|} \hline \mathbf{1} & \underline{2} & 7 & 4 \\ \hline \underline{5} & \mathbf{6} & 3 & 8 \\ \hline 9 & 10 & \underline{\mathbf{11}} & 12 \\ \hline 13 & 14 & 15 & \underline{\mathbf{16}} \\ \hline \end{array}$$

where bold entries correspond to the unique stable matching: $\mu_M(Fi) = Wi$ for all $i$. For sufficiently high $\delta < 1$, we show there is an equilibrium in weakly undominated strategies that implements the matching $\mu$: $\mu(F1) = W2, \mu(F2) = W1, \mu(F3) = W3$ and $\mu(F4) = W4$, corresponding to the underlined entries in the matrix.

Consider the following profile of strategies yielding the matching $\mu$. In period 1, firms $F2$ and $F4$ make an offer to worker $W1 = \mu(F2)$ and $W4 = \mu(F4)$, respectively. $W1$ and $W4$ accept these offers immediately, while workers $W2$ and $W3$ do not accept any offer, unless from their most preferred unmatched firm. In period 2, firms $F1$ and $F3$ make an offer to $W2 = \mu(F1)$ and $W3 = \mu(F3)$, respectively, who accept their offers.

With detectable deviations, if workers $W1$ and $W4$ do not receive an offer from any firm in period 1, they exit; otherwise, all workers reject any offer they receive, unless it is from their first-choice firm, and stay in the market until period 2. In period 2, if firms $F2$ and $F4$ have not matched with $\mu(F2)$ and $\mu(F4)$, respectively, $F3$ makes an offer to $W2$ instead of $W3 = \mu(F3)$ if possible. If $W2$ has already exited the market, $F3$ makes an offer to its most preferred unmatched worker, and exits if all workers have left the market. Furthermore, in period 2, any worker who does not receive offers exits immediately, and otherwise accepts his best offer. Any firm that gets rejected in period 1 exits the market in the beginning of period 2.

This profile constitutes an equilibrium in weakly undominated strategies that implements an unstable matching.[19] ||

The crux of the multiplicity in the example above is the limited restraint that elimination of weakly dominated strategies provides on off-equilibrium behavior. Specifically, $F3$

[19]When firms and workers can use weakly dominated strategies, there are even more equilibrium profiles and outcomes. For instance, it is an equilibrium for all agents to exit the market in period 1, resulting in no individual matches. Weak dominance rules this out, as it does not allow a worker to exit the market when he has an acceptable offer in hand.

can "punish" $F2$ for not making an offer to $W1 = \mu(F2)$ in period 1. Imposing subgame perfection would have little impact in our matching market game since individuals do not observe others' offers and reactions: the set of proper subgames is limited after period 1. Nonetheless, iterated elimination of weakly dominated strategies rules out strategies such as those driving Example 1 and, in fact, guarantees that the stable matching is the unique equilibrium outcome.

**Proposition 2** (Complete Information)**.** *For any complete-information market game, there exists a Nash equilibrium in weakly undominated strategies that yields the stable matching. Furthermore, the stable matching is the unique Nash equilibrium outcome surviving iterated elimination of weakly dominated strategies.*

When using strategies that survive iterated elimination of weakly dominated strategies, firms and workers that form top-top matches must be matched in period 1. Consider the top-top matches in the remaining submarket. Since the corresponding workers realize the top-top matches in the original market are formed in period 1, iterated elimination of weakly dominated strategies ensures that they accept their top-top matches in the remaining submarket. Therefore, the corresponding firms make those offers and are matched in period 1 as well. Continuing recursively, we get that iterated elimination of weakly dominated strategies guarantees that the unique stable matching of the market is implemented in one period.[20]

The above construction hinges on all agents being completely informed of the realized market, and hence able to compute the stable matching. Proposition 2 shows that a robust non-cooperative market game equilibrium generates the unique stable matching. In what follows, we show that this conclusion may change dramatically when agents exhibit incomplete information.

## 4 Decentralized Incomplete Information Economies

In economies with incomplete information, more than one market may realize. For participants to reach a stable outcome, sufficient information has to be transmitted to ensure

[20]Preference alignment plays a role. Without alignment, equilibria surviving iterated elimination of weakly dominated strategies may generate multiple outcomes, even with complete information.

that (i) firms make offers to workers who are their stable match partners in the realized market, and (ii) workers only accept offers from firms that are their stable match partners in that market.

There are three channels through which information flows in the decentralized market game. First, information is publicly transmitted when agents exit the market or form a match. Second, information is privately transmitted when workers receive offers from firms and workers respond to those offers (unless offers are accepted, in which case that information becomes public). The third component of information is time, which all market participants track.

## 4.1 The No-Discounting Case

Suppose there is no discounting, i.e., $\delta = 1$. Then, one way by which information can be transmitted in the market is through agents simply mimicking the DA algorithm. That is, firms make offers to workers according to their match utilities, and exit the market only when all workers have rejected them. Workers hold on to their best available offer, and accept an offer from any firm that generates the highest match utility among the set of available firms. These prescriptions are followed by all agents after any detectable deviations as well, and hold independently of the initial distribution $G$. We call these **DA strategies**.

**Proposition 3** (No Discounting)**.** *Suppose $\delta = 1$. The DA strategy profile constitutes a Bayesian Nash equilibrium in weakly undominated strategies and yields the stable matching.*

Intuitively, when agents use DA strategies, workers ultimately hold offers from their stable match partners. The top-top match property implies that, in every period, either a match is formed, or only agents on one side remain unmatched. Thus, the game terminates in finite time. When $\delta = 1$, the timing of matches is inconsequential to market participants and unilateral deviations cannot generate a preferable match. Nonetheless, as Example 1 already illustrated, uniqueness of equilibrium outcomes, or even the resulting matchings, is not guaranteed without further refinements.

### 4.2 Instability in Economies with Discounting

When $\delta < 1$, DA strategies are in general no longer incentive compatible. As an example, consider a complete information economy with two workers and two firms, for which $u^f_{ij} = u^w_{ij}$ for all $i, j$ and match utilities are given as follows:

$$U^f = U^w = \begin{array}{|c|c|}\hline \mathbf{4} & 1 \\ \hline 3 & \mathbf{2} \\ \hline \end{array}$$

The unique stable matching is highlighted in bold. Firm 2 knows that worker 2 is her unique stable match partner and, furthermore, that worker 2 would accept an offer from firm 2 immediately, as firm 2 is worker 2's first choice. Hence, it cannot be an equilibrium for firm 2 to first make an offer to worker 1 and delay her match. Firms may therefore be tempted not to make all offers in order of their preferences, but to instead concentrate on offers to plausible stable match partners. Similarly, workers may accept an offer from their highest plausible stable match partner, even if more preferred firms are still unmatched.

As it turns out, timing considerations are an inherent obstacle to stability in incomplete-information economies with frictions. The incentives to speed up matches may be so severe that no equilibrium generates the stable outcome in all markets in the economy's support.

**Example 2 (Manipulating Beliefs to Speed Up Matches)** Consider an economy with two firms $\{F1, F2\}$ and two workers $\{W1, W2\}$ with $u^f_{ij} = u^w_{ij}$ for each of 6 potential markets, described by the following match utilities (notation as before):

$$U_1 = \begin{array}{|c|c|}\hline \mathbf{3} & 6 \\ \hline 4 & \mathbf{7} \\ \hline \end{array}\; U_2 = \begin{array}{|c|c|}\hline 3 & \mathbf{6} \\ \hline \mathbf{4} & 5 \\ \hline \end{array}\; U_3 = \begin{array}{|c|c|}\hline \mathbf{3} & 2 \\ \hline 4 & \mathbf{8} \\ \hline \end{array}\; U_4 = \begin{array}{|c|c|}\hline \mathbf{3} & 2 \\ \hline 1 & \mathbf{7} \\ \hline \end{array}\; U_5 = \begin{array}{|c|c|}\hline \mathbf{9} & 6 \\ \hline 8 & \mathbf{5} \\ \hline \end{array}\; U_6 = \begin{array}{|c|c|}\hline 7 & \mathbf{3} \\ \hline \mathbf{8} & 5 \\ \hline \end{array}$$

We show that there are no equilibria in weakly undominated strategies that always implement the stable outcome, highlighted in bold.

$U_3$ and $U_4$ guarantee that $F1$ sometimes makes an offer to $W1$ in any equilibrium when $W1$'s match utilities are $(3, 4)$.[21] Similarly, $U_5$ and $U_6$ guarantee that $W2$ with match

[21] Indeed, in $U_4$, $W1$ accepts an offer from $F1$ immediately. Therefore, $F1$ has an incentive to make an offer to $W1$ in period 1 whenever $U_4$ is realized. However, $F1$ cannot distinguish between $U_3$ and $U_4$, so what are possible consequences of an offer to $W1$ in $U_3$? Given an offer from $F1$, $W1$ cannot exit (but he can reject the offer from $F1$). In $U_3$, it must be the case that $F2$ makes an offer to $W2$ that gets immediately accepted.

utilities $(6,5)$ will, in equilibrium, sometimes receive an offer in period 2, but not in period 1.[22]

From now on, we focus on $U_1$ and $U_2$. In particular, $F1$ observes $(3,6)$, $W1$ observes $(3,4)$, and they cannot distinguish between $U_1$ and $U_2$.

$W1$ and $W2$ always accept an offer from $F2$ immediately in $U_1$. Hence, for any $\delta < 1$, $F2$ must make an offer to $W2$ when $U_1$ is realized.

Assume that conditional on $F1$'s match utilities being $(3,6)$, the probability of $U_1$ is $p$ and that of $U_2$ is $1-p$.

**Suppose** $F1$ **makes an offer to** $W2$ **(in** $U_1$ **and** $U_2$**) in period** 1. Then, when $U_2$ prevails, $F2$, who is aware $U_2$ is realized, makes an offer to $W1$ in period 1, who accepts that offer. These strategies generate a payoff for $F1$ of $6(1-p)+3p\delta$.

Consider $F1$'s deviation to making an offer to $W1$ in period 1. Along the equilibrium path, $F1$ makes an offer to $W1$ with match utilities of $(3,4)$ only in $U_3$, when $F1$ is his stable match. Thus, $W1$ accepts an offer from $F1$ whenever $W1$'s match utilities are $(3,4)$ (and that is the only offer he observes). Hence, $W1$ accepts $F1$'s offer also in $U_1$. In $U_2$, the offer is rejected, and $F1$ matches with $W2$ in period 2 (as $W2$ does not leave the market in period 1 when observing match utilities $(6,5)$, see above), resulting in payoffs $6(1-p)\delta+3p$. This deviation is profitable when $p > 2/3$, independent of $\delta$.

Intuitively, $F1$ can be sure that, when approaching $W1$ in period 1, its offer gets accepted only when $W1$ is the stable match. The effect of such a deviation is therefore to *speed up* the creation of its match when $U_1$ is realized. The cost is the delay of a match with $W2$ in $U_2$. However, when $U_1$ is sufficiently more likely ex-ante (given $F1$'s private information), the benefits outweigh the costs.

**Suppose** $F1$ **makes an offer to** $W1$ **with probability** $q \in (0,1]$ **(in** $U_1$ **and** $U_2$**) in period** 1. This implies that $W1$ has to accept the offer from $F1$ with positive probability $m \in$

Therefore, in period 2, if $W1$ rejected the offer of $F1$, then $F1$ can remake that offer, in which case $W1$ has to accept it whenever using a weakly undominated strategy. Hence, in equilibrium, $F1$ best responds by making an offer to $W1$ whenever her match utilities are $(3,2)$.

[22] In $U_5$, $F1$ makes an offer to $W1$ who accepts immediately. In $U_6$, to guarantee a stable outcome, in period 1, $F2$ with utilities $(8,5)$ cannot make an offer to $W2$ and hence has to make an offer to $W1$. This implies that in $U_5$, $W2$ does not receive any offers in period 1 in equilibrium. However, $W2$ receives an offer from $F2$ in period 2. Therefore, $W2$ cannot exit the market when he receives no offer in period 1.

$(0,1]$.[23] In order for the market to always yield a stable outcome, it must be that $F2$ makes an offer to $W1$ with probability 1 in period 1 when $U_2$ prevails, and $W1$ accepts the offer from $F1$ whenever he receives it. Now, $F1$ has the same trade-off as before, and hence, $F1$ strictly prefers making an offer to $W1$, implying that $q = 1$. Can we induce $F2$ to make an offer to $W1$ with certainty? When $U_2$ prevails, an offer to $W1$ yields 4. An offer to $W2$ yields $5\delta$, which is bigger than 4 for $\delta > 4/5$. Hence, for large enough $\delta$, $F1$ making an offer to $W1$ with positive probability cannot be part of an equilibrium.

Suppose $F1$ simply delays making an offer and makes an offer to its most preferred available worker in period 2. This clearly cannot be part of an equilibrium, since $F1$ can profitably deviate by making an offer to $W2$ in period 1, which will be accepted with probability $1 - p$.

This example relies on $F1$'s incentive to make an offer to a worker ranked *below* her favorite unmatched plausible stable match partner. In particular, if there were an economy where $F1$ could be accepted by $W1$ when $W1$ was not its stable matching partner, this would incentivize $F1$ not to deviate for sufficiently high $\delta$. For instance, suppose the following market were added with positive probability (forming an economy identified by $U_1 - U_7$, each occurring with positive probability):

$$U_7 = \begin{array}{|c|c|} \hline 3 & 6 \\ \hline 2 & 1 \\ \hline \end{array}$$

$W1$ would immediately accept an offer from $F1$, penalizing speeding-up attempts in other states. ||

Example 2 illustrates a speeding-up motive that prevents the stable matching from being achieved. Agents' updating may also introduce incentives to manipulate beliefs in order to affect *who* an agent matches with, which the following example illustrates.

**Example 3 (Manipulating Beliefs to Improve Matches)** Consider an economy with three firms $\{F1, F2, F3\}$ and three workers $\{W1, W2, W3\}$ in which $u_{ij}^f = u_{ij}^w$ for each of 4 potential markets described by the following match utilities:

[23] Suppose $W1$ accepts $F1$ with probability 0 in period 1. Then $F1$'s payoff from an offer to $W1$ in period 1 is $3p\delta + 6(1-p)\delta$. However, an offer to $W2$ yields an expected payoff of $3p\delta + 6(1-p)$, and $F1$ would have a profitable deviation.

$$U_1 = \begin{array}{|c|c|c|}\hline \mathbf{5} & 4 & 3 \\ \hline 7 & \mathbf{10} & 2 \\ \hline 6 & 8 & \mathbf{9} \\ \hline \end{array} \quad U_2 = \begin{array}{|c|c|c|}\hline 5 & 9 & \mathbf{4} \\ \hline 7 & \mathbf{10} & 3 \\ \hline \mathbf{6} & 8 & 2 \\ \hline \end{array} \quad U_3 = \begin{array}{|c|c|c|}\hline \mathbf{5} & 4 & 3 \\ \hline 9 & 7 & \mathbf{10} \\ \hline 6 & \mathbf{8} & 2 \\ \hline \end{array} \quad U_4 = \begin{array}{|c|c|c|}\hline \mathbf{5} & 4 & 3 \\ \hline 2 & 6 & \mathbf{9} \\ \hline 1 & \mathbf{8} & 7 \\ \hline \end{array}$$

Suppose there exists an equilibrium in weakly undominated strategies that always implements the stable matching, highlighted in bold.

$U_3$ guarantees that, in such an equilibrium, $F3$ with match utilities $(6,8,2)$ always makes an offer to $W2$ in period 1.[24] Furthermore, $U_4$ guarantees that, in any such equilibrium, $F1$ with match utilities $(5,4,3)$ makes an offer to $W1$ with probability 1.[25] From here on, we concentrate on markets corresponding to $U_1$ and $U_2$.

**Suppose $F1$ makes an offer to $W3$ in $U_2$ in period 1 with certainty**. In this case, $W1$ receives an offer from $F1$ only when $F1$ is the stable match partner (as in $U_2$, $F3$ makes an offer to $W2$ in period 1 and to $W1$ only in period 2). Thus, in equilibrium, $W1$ will accept an offer from $F1$ in period 1 if it is the only offer he receives. This provides strict incentives for $F1$ to make an offer to $W1$ even in $U_2$, resulting in an outcome that is not stable.

**Suppose $F1$ makes an offer to $W1$ in $U_2$ with probability $q \in (0,1]$**. When $F1$ makes an offer to $W3$, $F1$ receives a payoff of 4, as $W3$ accepts that offer immediately. In order for the market to always yield a stable outcome, it has to be the case that $W1$ never accepts an offer from $F1$ in period 1 when he receives only that offer. Therefore, the expected payoff of $F1$ from making an offer to $W1$ is $4\delta < 4$, in contradiction to $F1$ playing a best response.

In the example, $W1$ cannot distinguish $U_1$ from $U_2$. Hence, his set of plausible stable match partners at $t = 1$ is $\{F1, F3\}$ when either $U_1$ or $U_2$ is realized. At the heart of the difficulty of achieving the stable matching in equilibrium is the fact that $W1$ cannot be certain whether he will receive his best offer in period 1 or in period 2. $W1$ tries to infer that from offers received in period 1. However, offers can then be manipulated. The

[24]Such an offer is immediately accepted when $U_3$ prevails. Furthermore, if $F2$ makes an offer to $W1$ with some probability $p > 0$ in period 1, then when $U_3$ prevails, since $W1$ can infer the stable match partner is $F1$ and generates a match utility of 5, $W1$ accepts that offer, yielding an unstable outcome.

[25]When $U_4$ prevails, $W1$ accepts an offer from $F1$ immediately. When $U_3$ prevails, $F2$ matches with $W3$ and $F3$ with $W2$ in period 1 (see above). Hence, when $U_3$ prevails, $W1$ matches with $F1$ in period 2 at the latest, so $F1$ does not lose anything from making an offer to $W1$ in period 1.

example illustrates the potential for manipulation of offers when information regarding the set of potential stable match partners is transmitted by the mere timing of an offer (or the acceptance of an offer). This form of information transmission needs to be restricted to allow for equilibria that yield the stable match. As in Example 2 above, the addition of markets can help prevent this problem by limiting learning that occurs through the timing of events per se. For instance, suppose we add another market to the economy of Example 3, described by match utilities $U_5$ as follows:

$U_5 =$

| | | |
|---|---|---|
| 5 | 4 | **3** |
| **7** | 10 | 2 |
| 6 | **11** | 5 |

with the likelihood of each market in the economy being arbitrary, but strictly positive. In this augmented economy, for sufficiently high discount factors, $F2$ makes an offer to $W2$ since she cannot tell whether $U_1$ or $U_5$ govern match utilities. Certainly, $F3$ makes an offer to $W2$ as well since he is her stable match partner and her favorite. It follows that, when $U_5$ prevails, $W1$ receives only one offer from $F1$ at $t = 1$, which he should no longer accept. This breaks the incentive of $F1$ to deviate as described in the example. ||

### 4.3 Rich Economies and Stable Equilibrium Outcomes

Examples 2 and 3 suggest that obstacles to stability emerge when learning occurs via the timing of events. At one extreme, when all participants are informed of the realized market, there is no learning to be had and stable outcomes can be implemented in equilibrium. As both examples illustrate, the addition of certain markets limits the scope for manipulation as well. We now show that, indeed, when there are many possible markets in the economy, learning is restrained and stability can be achieved in equilibrium.

We introduce *rich economies*, where the support of $G$ contains markets representing all possible constellations of aligned preferences. In rich economies, an agent's utility type provides limited information regarding other agents' utility types. Furthermore, all matchings are possibly stable given an agent's private information at the outset of the market game. As we show, this restricts any agent's ability to profitably manipulate others' beliefs.

Formally, a **rich economy** is an economy in which the distribution $G$ over markets has full support over all possible aligned preferences represented with payoffs from $\Pi \subseteq \mathbb{R}_{>0}$ with $|\Pi| \geq \max\{F, W\} + 1$.[26]

Richness ensures that every possible preference constellation associated with an aligned market has a positive probability. In particular, it drastically limits what agents can infer about others' preferences from their own realized match utilities. We stress that richness imposes no restrictions on the likelihood of any market in the economy, which can be arbitrarily small.

**Proposition 4** (Stable Implementation in Decentralized Economies)**.** *Suppose the economy is rich. For sufficiently high $\delta$, there exists a Bayesian Nash equilibrium in weakly undominated strategies of the decentralized market game that implements the unique stable matching in each supported market.*

We describe the intuition for the proof of Proposition 4 in Subsection 4.4 below. While we present Proposition 4 for rich economies, in the Appendix, we provide more general conditions on the support of the distribution $G$ that allow for the implementation of stable outcomes. The conditions formalize the restrictions that, under a desirable class of strategy profiles that culminate in stable matchings, the timing of events does not reveal information in and of itself. For instance, full support on all aligned markets in which preferences are "assortative" on one side—say, where all firms rank workers in the same way—also ensures implementation of stable outcomes.

Proposition 4 suggests that stability may be reached even with incomplete information and time frictions when economies are rich. It does not, however, imply uniqueness of these outcomes. In fact, uniqueness in such economies is much harder to guarantee and is more heavily tied with the environments' details, even when market participants are arbitrarily patient. Indeed, even when the economy is rich, suppose one market is far more likely than others and consider the following strategy profile. Firms make an offer to their stable match partner in that market. Workers accept their best offer that is superior to

[26]The requirement that $|\Pi| \geq \max\{F, W\} + 1$ is needed since we assume agents are never indifferent. Indeed, suppose $\max\{F, W\} = F$. For each worker, we need to specify the match utility generated by all $F$ firms, as well as the match utility from remaining unmatched, necessitating $F + 1$ values.

their stable match partner in that market, and exit otherwise. Similarly, any firm that is not accepted at $t = 1$ exits immediately. When the probability of the most likely market is sufficiently high, this profile constitutes an equilibrium in weakly undominated strategies. In fact, the profile survives iterated elimination of weakly dominated strategies. It allows for the stable matching to be implemented in the most likely market, but not in all others. Thus, the uniqueness apparent under mild refinements from Proposition 2 when information is complete does not carry over to environments with incomplete information.

### 4.4 Implementing Stable Matchings—Proof Intuition

Before approaching the proof of Proposition 4, we analyze some minimal conditions strategies have to satisfy in a centralized firm-proposing DA mechanism in order to guarantee a stable outcome. This allows us to ignore incentive compatibility hurdles that are due to interim learning. In the proof of Proposition 4, we illustrate how richness guarantees that at least some of these strategy profiles, translated into our dynamic game, are incentive compatible for all participants.

In the centralized setting, if there is any hope of achieving the complete-information stable matching for any market realization, agents must declare plausible stable match partners acceptable. Furthermore, consider the firms, for example. Since the firm-proposing DA mechanism achieves the firm-optimal stable matching for submitted preferences, permuting the ranking of agents that are less preferred than a firm's stable match partner would not change the resulting matching. However, it is crucial that agents ranked above *any* plausible stable match partner are, in fact, preferred to that stable match partner. These restrictions suggest a class of strategies.

Formally, for any $l \in \mathcal{F} \cup \mathcal{W}$ of type $\alpha \in \{f, w\}$, let $S_l^\alpha\left(u_l^\alpha(\cdot)\right)$ denote agent $l$'s plausible stable match partners at the outset, using only the information in $l$'s own match utilities.[27] When agent $l$ submits utilities $v$ corresponding to preferences $P(v)$, let $A(v) = \{k : kP(v)\varnothing\} \cup \{\varnothing\}$, where $A(v) \subseteq \mathcal{W} \cup \{\varnothing\}$ if $l \in \mathcal{F}$ and $A(v) \subseteq \mathcal{F} \cup \{\varnothing\}$ if $l \in \mathcal{W}$. That is, $A(v)$ is the union of $l$'s acceptable match partners given $P(v)$ and the option of staying unmatched, namely the "weakly acceptable" partners of $l$ under $P(v)$.

[27] If $\alpha = f$ (and $l \in \mathcal{F}$), then $u_l^\alpha(k) = u_{lk}^f$; otherwise, $u_l^\alpha(k) = u_{kl}^w$.

**Definition** (Reduced DA Strategies)**.** In an economy $\mathcal{E}$, suppose agent $l \in \mathcal{F} \cup \mathcal{W}$ of type $\alpha \in \{f, w\}$ submits utilities $v$ corresponding to preferences $P(v)$. Then $v$ is a reduced DA strategy if:

1. $S_l^\alpha\left(u_l^\alpha(\cdot)\right) \subseteq A(v)$.

2. For each $k, m \in S_l^\alpha\left(u_l^\alpha(\cdot)\right)$, $kP(v)m \Leftrightarrow u_l^\alpha(k) > u_l^\alpha(m)$.

3. For each $k \in A(v) \backslash S_l^\alpha\left(u_l^\alpha(\cdot)\right)$ and each $m \in S_l^\alpha\left(u_l^\alpha(\cdot)\right)$, $kP(v)m \Rightarrow u_l^\alpha(k) > u_l^\alpha(m)$.

The first restriction requires that any plausible stable match partner—using the agent's private information on match utilities—is declared acceptable. The second restriction states that plausible stable match partners are ranked truthfully. The third restriction is perhaps more subtle. To glean intuition, suppose $k$ is a "weakly acceptable" match partner for a worker $l$ who is not a plausible stable match partner. Suppose further that our worker $l$ ranks $k$ above a stable match partner $m$, although he prefers $m$. Since $m$ is a plausible stable match partner, in some plausible market, agent $l$ receives an offer from $m$. Under the declared preferences, such an offer may end up being rejected in favor of an offer from $k$, which is less preferred and not necessarily part of a stable matching. The third restriction rules out such reports, with similar intuition holding for firms.

We now show the sense by which reduced DA strategies place minimal requirements for securing stability in centralized markets. Note that the conditions of reduced DA strategies depend only on the *support* of stable matchings, and not on their precise likelihood of occurrence. We consider economies with a finite number of potential markets: strategy profiles that generate the stable matching need to be robust to how these markets are distributed. We say an agent $l$ uses a **rule** if, for any economy $\mathcal{E}$ containing agent $l$, the agent uses a strategy that depends only on the set of market participants, the agent's realized match utilities, and the set of potential stable match partners. Agent $l$ uses a **reduced DA rule** if the utilized strategy is a reduced DA strategy.[28]

[28] Reduced DA *rules* impose restrictions on the details of the economy agents can utilize in their strategies. In particular, suppose $\mathcal{E}_1$ and $\mathcal{E}_2$ are two economies with the same set of firms and workers, both containing a market $M = (U, \mathcal{F}, \mathcal{W})$ such that, for some agent $l \in \mathcal{F} \cup \mathcal{W}$ of type $\alpha \in \{f, w\}$, the set of a priori stable matches $S_l^\alpha(u_l^\alpha(\cdot))$ is identical when $M$ is realized in either economy. Then, if agent $l$ uses a reduced DA rule, he or she must use the *same* reduced DA strategy in both $\mathcal{E}_1$ and $\mathcal{E}_2$, whenever observing $u_l^\alpha(\cdot)$.

**Lemma 2** (Stable Implementation in Centralized Economies)**.**

1. *If all agents use a reduced DA rule then, for any economy $\mathcal{E}$, the outcome produced by the DA mechanism is the stable matching.*

2. *Suppose an agent $a \in \mathcal{F} \cup \mathcal{W}$ uses a rule that is not a reduced DA rule. Then there exists an economy $\mathcal{E}$ for which the DA mechanism produces an outcome that is not stable in some market realization.*

3. *Given an economy $\mathcal{E}$, all agents using reduced DA strategies constitutes a Bayesian Nash equilibrium of the game associated with the DA mechanism.*

Lemma 2 illustrates the effectiveness of reduced DA rules in generating stable outcomes. Part 2 of the lemma highlights the necessity of the conditions imposed by reduced DA strategies for implementing stable outcomes in any economy. Part 3 is a strengthening of the second part of Lemma 1. It highlights, again, the fact that in centralized matching markets, implementation of stable outcomes can be done through equilibrium, even with incomplete information.

When moving from a centralized to a decentralized market, we need to translate reduced DA strategies to dynamic prescriptions. For firms, the translation is straightforward. Whenever firm $i$ submits a reduced DA strategy $v$, it acts as follows in the decentralized market game. In each period in which firm $i$ is not matched and does not have an offer held by a worker, firm $i$ makes an offer to its most preferred unmatched worker who has not rejected firm $i$ yet according to $v$ (where ties are broken according to the same rules determining $P(v)$ in the centralized setting, depending on the index of the match partner and in favor of matching). When firm $i$ gets rejected by the last acceptable worker (according to $v$), firm $i$ exits the market.

For workers, there are two aspects of strategies that are important. The first is when to start accepting offers, the second is which offer to accept. In terms of the latter, the use of weakly undominated strategies implies that when a worker accepts an offer, it must be his favorite available. Thus, each worker has to rank *all* firms in the “right” order. However, in a decentralized market, a worker may accept an offer even if it is not from his most preferred unmatched firm. The translation of a reduced DA strategy $v$ to a decentralized market is captured by "threshold firms." At each point in time, a worker accepts the best available offer that he likes at least as much as his most preferred unmatched plausible stable match partner, ranked according to $v$: the "threshold firm." So, if the reduced DA strategy $v$ only ranks potential stable match partners, the worker accepts an offer as soon as he receives an offer he prefers at least as much as the most preferred unmatched potential stable match partner. In order to mimic the operation of reduced DA strategies within the firm-proposing DA mechanism, we require workers to hold their best available offer as long as the offer is at least as good as their lowest potential stable match, and reject all other offers.

When moving from a centralized to a decentralized market, agents may also condition their actions on the history of play. Specifically, in decentralized markets, given the strategy profile of all other agents, each firm or worker can recalculate and possibly refine their set of plausible stable match partners over time. **Decentralized reduced DA strategies** are therefore strategies that can be derived as above when allowing agents to submit a new reduced DA strategy every period. This allows agents to take information they have accumulated into account. On the firm side, this implies that a firm makes an offer to workers that are at least as good as her most preferred unmatched plausible stable partner who has not rejected her yet. On the worker side, this has three implications. First, the best offer that is weakly preferred to a worker's least preferred plausible stable match is not rejected. That is, a worker never rejects an offer that could potentially be superior to his stable match partner. Second, upon receiving an offer at least as good as his (strategically specified) most preferred unmatched firm, a worker immediately accepts that offer. Last, in analogy to the operation of the DA algorithm, the worker rejects all offers that are not as good as the least preferred plausible stable partner, and only holds one offer.

In the proof of Proposition 4, we show that beliefs cannot be beneficially manipulated through deviations from decentralized reduced DA strategy profiles when $\delta$ is high. The key intuition is the following. Through deviating, an agent attempts to manipulate beliefs at the possible cost of relinquishing the stable matching. For example, a firm makes an offer to a lower-ranked worker or a worker rejects its best-held offer. Since all markets have full support, with positive probability, such manipulations can cause an agent to miss out on the best match partner and experience a strict utility loss. Furthermore, the no-cycle property of aligned markets ensures that rejecting a good offer cannot generate a superior offer. In principle, deviations may expedite the time at which an agent is matched. However, for sufficiently high $\delta$, the strictly positive loss from losing the most-preferred match overwhelms the gain due to timing.

### 4.5 Duration of Decentralized Interactions

With richness, Proposition 4 guarantees the implementation of stable matchings even in the presence of information and time frictions. One potential concern, however, pertains to the time it takes markets to stabilize. In practice, prolonged durations of instability may entail substantial efficiency losses: the use of unemployment benefits, limited learning-by-doing of the labor force, and so on.

The longest any realized market might take to reach a stable outcome is given by $\min\{F, W\}$. This is due to our alignment assumption. In each period, at least one top-top match is present. Therefore, at least one firm and one worker exit each period until one side of the market is exhausted. Nonetheless, in principle, stable outcomes can be achieved more rapidly for "most" conceivable markets. In order to get a sense of these durations of decentralized interactions, we consider markets in which firms' and workers' preferences are determined uniformly at random from all possible aligned preference profiles. We run 10,000 simulations for each market size—the number of firms and the (equal) number of workers. We assume agents follow the decentralized reduced deferred acceptance strategies.[29] Figure 1 illustrates the average number of periods markets take to reach the stable outcome.

As can be seen, in our simulated markets, the number of periods required to reach stability is far smaller than our upper bound. For instance, for a market with 400 participants (200 firms and 200 workers), the average time to reach stability is shorter than 9 periods.

[29] It is straightforward to see that, in rich economies, all decentralized reduced DA strategy profiles generate the same duration of interactions till the market game ends.

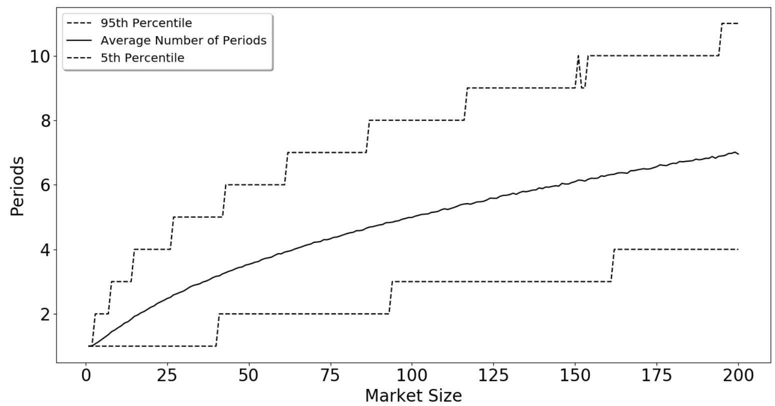


Figure 1: Duration of Decentralized Interactions across Market Sizes

## 5 Conclusions

We analyze a decentralized market game in which firms make offers that workers react to, allowing for incomplete information and time frictions. We identify when there exists an equilibrium in weakly undominated strategies that yields the complete-information stable matching. This is the case with complete information, when the economy consists of a single market. It is also the case when there are no time frictions. With both uncertainty and time frictions, the economy needs to be sufficiently rich and discounting has to be insubstantial enough for there to be an equilibrium that yields stability. For such settings, our analysis provides a non-cooperative foundation for the cooperative stability notion.

Taken together, our results indicate that when studying markets, it is crucial to understand characteristics that go beyond the identification of market participants and their preferences. The information available to participants and the plausibility of frictions both play an important role in predicting outcomes. This, in turn, implies that channels by which information can be transmitted among market participants can be a critical element of market design.

## 6 Appendix – Proofs

The following notation will be useful for several of our proofs. Suppose preferences are aligned with ordinal potential $\Phi = \left(\Phi_{ij}\right)$.

Let

$$M^{(1)} = \{(i,j) \in \mathcal{F} \times \mathcal{W} | (i,j) \in \arg\max_{(i,j)\in\mathcal{F}\times\mathcal{W}} \Phi_{ij}\}.$$

Define

$$\begin{aligned} \mathcal{F}_M^{(1)} &= \{j | \exists i \in \mathcal{W} \text{ s.t. } (i,j) \in M^{(1)}\}, \\ \mathcal{W}_M^{(1)} &= \{i | \exists j \in \mathcal{F} \text{ s.t. } (i,j) \in M^{(1)}\}. \end{aligned}$$

Denote now

$$\mathcal{F}^{(2)} = \mathcal{F} \setminus \mathcal{F}_M^{(1)}, \mathcal{W}^{(2)} = \mathcal{W} \setminus \mathcal{W}_M^{(1)}.$$

The submarket corresponding to $\mathcal{F}^{(2)}$ and $\mathcal{W}^{(2)}$ has aligned preferences and $\Phi$ restricted to those firms and workers serves as an ordinal potential. We can replicate the construction above recursively and define, for any $k$,

$$\begin{aligned} M^{(k)} &= \{(i,j) \in \mathcal{F}^{(k)} \times \mathcal{W}^{(k)} | (i,j) \in \arg\max_{(i,j)\in\mathcal{F}^{(k)}\times\mathcal{W}^{(k)}} \Phi_{ij}\}, \\ \mathcal{F}_M^{(k)} &= \{j : \exists i \in \mathcal{W} \text{ s.t. } (i,j) \in M^{(k)}\}, \\ \mathcal{W}_M^{(k)} &= \{i : \exists j \in \mathcal{F} \text{ s.t. } (i,j) \in M^{(k)}\}, \\ \mathcal{F}^{(k+1)} &= \mathcal{F} \setminus \mathcal{F}_M^{(k)}, \text{ and } \mathcal{W}^{(k+1)} = \mathcal{W} \setminus \mathcal{W}_M^{(k)}. \end{aligned}$$

The unique stable matching $\mu_M$ can be identified using this recursion: for each $k$, whenever $(i,j) \in M^{(k)}$, $\mu_M(i) = j$. Any unassigned remaining agents are unmatched under $\mu_M$.

**Proof of Proposition 1**

1. A rejection cycle is equivalent to the existence of a weak improvement cycle in the two-player game with payoff matrix $((u_{ij}^w, u_{ij}^f))_{i,j}$. Our claim then follows directly from Voorneveld and Norde (1997).

2. Suppose preferences are aligned with ordinal potential $\Phi = \left(\Phi_{ij}\right)$. For any submarket $(\tilde{\mathcal{F}}, \tilde{\mathcal{W}}, \tilde{U})$, any pair $(i,j) \in \arg\max_{(i,j)\in\tilde{\mathcal{F}}\times\tilde{\mathcal{W}}} \Phi_{ij}$ constitutes a top-top match, and the top-top match property holds.

In addition, suppose $\mu$ is a matching different than the unique stable matching $\mu_M$. Reconstruct $\mu_M$ as above. Consider the smallest $k$ such that there is a pair $(i, j)$ that is matched under $\mu_M$ and not under $\mu$. In that case, $(i, j)$ blocks $\mu$ and $(i, j) \in M^{(k)}$ form a top-top match in the submarket corresponding to $\mathcal{F}^{(k)}$, $\mathcal{W}^{(k)}$. As this is the first discrepancy between $\mu_M$ and $\mu$ in the iterative process, the match partner $\mu(i)$ of $i$ and $\mu(j)$ of $j$ are part of the remaining set of firms and workers and hence inferior to $\mu_M(i) = j$ and $\mu_M(j) = i$, respectively. That is, the stable blocking property holds. ■

**Proof of Proposition 2**

Consider any Nash equilibrium that survives iterated elimination of weakly dominated strategies. Workers using weakly undominated strategies assures that, at any stage, a worker who receives an offer from their most preferred available firm accepts that offer immediately. At period 1, any firm $i \in \mathcal{F}_M^{(1)}$ must make an offer to $\mu_M(i)$, who will accept immediately. Any worker in $\mathcal{W}_M^{(2)}$ should therefore accept an offer from firms in $\mathcal{F}_M^{(2)}$ immediately. It follows that each firm $i \in \mathcal{F}_M^{(2)}$ must make an offer to $\mu_M(i)$ in period 1 as well. Continuing recursively, we get that all firms that are matched under $\mu_M$ must make offers that get accepted immediately to their corresponding match partner under $\mu_M$. Thus, firms or workers that are not matched under $\mu_M$ must exit the market in period 1 when $\delta < 1$ or they must exit in some period $k$ when $\delta = 1$. In particular, any profile surviving iterated elimination of weakly dominated strategies generates the matches prescribed by $\mu_M$ in period 1. ■

**Proof of Proposition 3**

Since preferences are aligned and everyone is acceptable to everyone, in every period with unmatched agents, there is either a top-top match that is formed, or only agents on one side of the market are unmatched. Thus, DA strategies generate a market matching in finite time. From the convergence of the Gale-Shapley algorithm to a stable matching, it follows that DA strategies yield the stable matching. We now show that DA strategies constitute a Bayesian Nash equilibrium.

Workers can deviate in two ways. First, a worker $j$ can reject an offer from firm $i$ instead of holding it. From the no-cycle property, such a rejection cannot launch a chain generating a superior offer for $j$. In addition, if $i$ is a plausible stable match partner, such a rejection may lead $j$ to forgo his best offer in some market. Such a deviation could therefore be profitable only if it makes the worker sufficiently better off in some market realization in which firm $i$ is not his stable match partner. However, in any market, it cannot be that $i$ is strictly better than $j$'s stable match partner, as then $j$ should never receive an offer from $i$ (indeed, by the construction of DA strategies, firm $i$ and worker $j$ would form a blocking pair to the stable matching). The second potential deviation of a worker is the acceptance of an offer that is not from his most preferred unmatched firm. However, workers are made better off over time, as they receive new offers. Therefore, accepting an offer early cannot be a profitable deviation when $\delta = 1$.

Consider now the firms. Suppose firm $i$ deviates and makes an offer to worker $j$ who is not the most preferred worker among workers who have not rejected $i$ yet. Since $\delta = 1$, if there is a market in which $i$ strictly benefits from this deviation, it must be the case that $i$ ends up matching with a strictly preferable worker. Suppose the resulting market matching (assuming all other agents follow the DA strategies) is $\mu$. The matching $\mu$ has the property that the set of firms $F'$, who prefer this match to the stable match $\mu_M$, is non-empty, as it contains at least firm $i$. By the Blocking Lemma, there exists a blocking pair $(i', j')$ with $i'$ not in $F'$ such that $j'$ is matched in $\mu$ to a firm in $F'$ Roth and Sotomayor (1992). Since $i'$ and $j'$ follow DA strategies, $i'$ must have made an offer to $j'$, a contradiction. ■

**Proof of Lemma 2**

**1.** Assume all agents use a reduced DA rule and suppose $\mathcal{E}$ is an economy with a market realization in which the outcome $\mu'$ is different than the stable matching $\mu_M$. By Proposition 1, there exists a pair $(i', j')$ that blocks $\mu$ with $\mu_M(i') = j'$. This implies that $u^f_{i'j'} > u^f_{i'\mu(i')}$ and $j' \in \mathcal{S}^f_{i'}(u^f_{i'\cdot})$. Hence, if firm $i'$ uses a reduced DA strategy, $i'$ must rank $j'$ above $\mu(j')$. For worker $j'$, since $i' \in \mathcal{S}^w_{j'}\left(u^w_{\cdot j'}\right)$, $j'$ cannot be matched with someone other than $i'$ unless, within the DA algorithm, he receives a better offer. However, $u^w_{i'j'} > u^w_{\mu(j')j'}$ implies that $j'$ does not reject $i''$s offer, in contradiction.

**2.** Suppose $a \in \mathcal{F} \cup \mathcal{W}$ is an agent who does not use a DA rule. That is, agent $a$ ranks some agent (including potentially the possibility of staying unmatched, $\varnothing$) as preferred to a plausible stable match partner when match utilities prescribe otherwise. Whenever there is only one worker or only one firm, the claim follows trivially. Assume then that

$|\mathcal{F}|, |\mathcal{W}| \geq 2$.

Certainly, if $a$ ranks a plausible stable match partner as unacceptable, then whenever the market in which the unique stable matching entails that individual match for $a$, the centralized outcome is unstable.

Suppose that $a \in \mathcal{W}$ ranks a plausible stable match partner $i$ below a firm $i'$, who is not a plausible stable match partner when observing $u^w_{\cdot a}$, and the set of plausible stable matches is $S$. Assume $u^w_{ia} > u^w_{i'a}$. Let $j \in \mathcal{W}$ be another worker (other than $a$).

Consider an economy in which there are three markets characterized by match utilities $U, \tilde{U}$, and $\hat{U}$ in which, in the corresponding stable matchings, all agents $A$ other than $i, i', a$, and $j$ are prescribed to be matched to agents in $A$ or remain unmatched. It therefore suffices to focus on match utilities corresponding to agents $\{i, i', a, j\}$.

We construct $U$ and $\tilde{U}$ so that they satisfy the following:[30]

**a.** Firm $i'$ cannot distinguish between the two markets, while all other agents can.

**b**. Firm $i'$ prefers worker $a$ to worker $j$ in both markets.

**c**. Under $U$, $i$ and $a$, and $i'$ and $j$, are part of the stable matching, while under $\tilde{U}$, $i'$ and $a$ are part of the stable matching.

**d**. Under $\tilde{U}$, $i'$ is both $a$'s and $j$'s most preferred firm.

$\hat{U}$ is such that $\hat{u}^w_{\cdot j} = \tilde{u}^w_{\cdot j}$, so that worker $j$ cannot distinguish $\tilde{U}$ from $\hat{U}$, and $j$ is the most preferred worker for $i'$.

Each of the remaining markets in the economy is one in which $a$'s match utilities are given by $u^w_{\cdot a}$ and the stable matching is an element $i'' \in S \setminus \{i\}$.

If the stable matching is achieved under $\hat{U}$, worker $j$ must rank firm $i'$ as acceptable when observing $\hat{u}^w_{\cdot j} = \tilde{u}^w_{\cdot j}$. Therefore, if the stable matching is achieved under $\tilde{U}$, it must be

[30]These conditions are consistent with alignment. Indeed, assuming, without restriction, that $u^w_{kl}, \tilde{u}^f_{kl} > 2$ for any $k \in \{\varnothing, a, j\}$ and $l \in \{\varnothing, i, i'\}$, the reader can think of the following manifestation of $U, \tilde{U}$ in which we summarize preferences through the following two matrixes, where the first number in each entry corresponds to the firm's preference and the second number to the appropriate worker:

$U$:

| | $a$ | $j$ |
|---|---|---|
| $i$ | $u^w_{ia}, u^w_{ia}$ | $u^w_{ia} - 1, u^w_{ia} - 1$ |
| $i'$ | $u^w_{i'a}, u^w_{i'a}$ | $u^w_{i'a} - 1, u^w_{i'a} - 1$ |

$\tilde{U}$:

| | $a$ | $j$ |
|---|---|---|
| $i$ | $\tilde{u}^f_{i\varnothing} - 1, \tilde{u}^w_{\varnothing a} - 1$ | $\tilde{u}^f_{i\varnothing} - 2, \tilde{u}^w_{\varnothing j} - 1$ |
| $i'$ | $u^w_{i'a}, \tilde{u}^w_{\varnothing a} + 1$ | $u^w_{i'a} - 1, \tilde{u}^w_{\varnothing j} + 1$ |

.

the case that $i'$ ranks $a$ higher than $j$ (and acceptable) when observing $u^f_{i'\cdot}$. But then, under $U$, it cannot be the case that the stable matching is established. Indeed, the centralized mechanism generates a stable matching for the submitted preference rankings, and $i'$ and $a$ would form a blocking pair.

A similar construction can be presented if $a \in \mathcal{W}$ ranks a plausible stable match partner $i$ below a less preferred plausible stable match partner $i'$ when observing $u^w_{\cdot a}$ and the set of plausible stable matches is $S$. Furthermore, analogous constructions follow when agent $a$ is a firm that does not follow a reduced DA rule.

**3.** Assume all agents follow a profile $v$ of reduced DA strategies. For firms, since truthful revelation is a weakly dominant strategy in the firm-proposing DA algorithm, which the centralized market emulates, no deviation can be strictly profitable.

Suppose a worker $j$ has a strictly profitable deviation to a strategy $\sigma_j$. If $\sigma_j$ is also a reduced DA strategy, then by part 1 above, the outcome is unchanged, in contradiction. Suppose, then, that $\sigma_j$ yields a matching $\mu$ such that $u^w_{\mu(j)j} > u^w_{\mu_M(j)j}$. Then, by Proposition 1, there exists a pair $(i', j')$ that blocks $\mu$ such that $\mu_M(i') = j'$. First, it is clear that $j' \neq j$ since $j$ strictly prefers $\mu$ to $\mu_M$. Since both $i'$ and $j'$ submit reduced DA strategies, $i'$ must rank $j'$ above $\mu(i')$. Hence, it must be that $j'$ rejects $i'$ through the centralized mechanism, contradicting the fact that $(i', j')$ is a blocking pair under $\mu$. ■

Before providing the proof for Proposition 4, we introduce economies that are *generalized rich*. The rich economies described in the body of the text are, as we show later, generalized rich. There are two requirements for generalized richness that limit learning in a way that guarantees reduced DA strategies constitute an equilibrium. Generalized richness identifies a broader class of settings in which (complete-information) stable outcomes can be implemented through equilibrium of our decentralized market game.

In what follows, we will use $\mu_M(U)$ to denote the unique stable matching in a realized (aligned) market with match utilities $U$.

Some notation will be useful. At any period $t$, let $M_t \subseteq (\mathcal{F} \cup \varnothing) \times (\mathcal{W} \cup \varnothing)$ denote the matches formed at time $t$, including firms and workers who leave unmatched, and let the set of agents who exited the market up to, but excluding, period $t$ be

$$\mathcal{X}^t \equiv \{j \mid \exists k \text{ s.t. } (j,k) \in M_\tau \text{ for some } \tau < t\} \cup \{i \mid \exists l \text{ s.t. } (l,i) \in M_\tau \text{ for some } \tau < t\}.$$

Let $M_t^F \subseteq \mathcal{F} \times \varnothing$ be the set of firms who leave the market in the first stage of period $t$.

At the beginning of period $t$, each active firm $i$ observes a history that consists of the (timed) offers the firm made, the responses of workers to those offers, denoted by $r$ for rejection and $h$ for holding (we denote an offer to no worker as an offer to $\varnothing$ that is immediately rejected), and the (timed) set of agents that have left the market:[31]

$$h_{t,i}^f \in ((\mathcal{W} \cup \varnothing) \times \{r, h\})^{t-1} \times \prod_{\tau=0}^{t-1} M_\tau.$$

Each unmatched worker acts in the interim stage of every period $t$ and observes a history that consists of all (timed) offers he received, including those at time $t$, a (timed) sequence of offers he has held, and the (timed) set of agents that have left the market up to and including time $t$:

$$h_{t,j}^w \in \left(2^{\mathcal{F}}\right)^t \times \left(2^{\mathcal{F}}\right)^t \times \prod_{\tau=0}^{t-1} M_\tau \times M_t^F.$$

For a given prior distribution $G$ over utility realizations, for any private information $u_l^\alpha(\cdot)$ of agent $l$ with $\alpha \in \{f, w\}$ regarding the realized market, let $G(u_l^\alpha(\cdot))$ denote the posterior distribution over utility realizations. Let $\mathcal{S}_l^\alpha(u_l^\alpha(\cdot)) = \left\{\mu_M(U)(l) \mid U \in \text{supp } G(u_l^\alpha(\cdot))\right\}$ denote the set of all ex-ante plausible stable match partners of agent $l$. That is, agents that could conceivably be part of a stable matching, under the distribution over market match utilities updated by the private information $u_l^\alpha(\cdot)$. Analogously, given the strategies played by all agents, $\mathcal{S}_l^\alpha(u_l^\alpha(\cdot), h_{t,l}^\alpha)$ denotes the set of all interim potential stable match partners given agent $l$'s available information at $t$.

**Assumption 1** *Suppose all agents use decentralized reduced DA strategies. Consider any firm $i$ and market realization with match utilities $U$. At each period $t$, for all available workers $j$ that are worse than firm $i$'s most-preferred potential stable match partner that has not rejected the firm yet, either*

1. *there exists $\tilde{U} \in$ supp $G\left(u_{i\cdot}^f, h_{t,i}^f\right)$ such that $\tilde{u}_{ij}^w > \tilde{u}_{\mu_M(\tilde{U})(j)j}^w$, or*

[31] An offer of firm $i$ to worker $j$ that is held from period $t$ to $t'$ is recorded as an offer made in periods $t, t+1, \ldots, t'$ that is held by the worker in each of these periods.

2. *for all* $\tilde{U} \in supp\ G\left(u^f_{i\cdot}, h^f_{t,i}\right)$, $\tilde{u}^w_{ij} < \min\{\tilde{u}^w_{kj} : k \in \mathcal{S}^w_j\left(\tilde{u}^w_{\cdot j}, h^w_{t,j}\right)\}$.

Assumption 1 ensures that when all market participants follow decentralized reduced DA strategies, a firm has no incentive to make an offer to a worker who is ranked below her favorite unmatched plausible stable partner that has not rejected her yet. If the firm makes such an offer, Assumption 1 guarantees the firm runs the risk of one of two eventualities. The first is that the firm might have her offer held or accepted, as it is better than the stable match of the worker in the realized market. The second occurs if the worker uses a decentralized reduced DA strategy that specifies firms that are less preferred than all plausible stable matches as unacceptable. In that case, the firm will be rejected immediately. Making an offer then has no benefit.

Denote by $\tilde{W}^t_i$ the set of workers who have not rejected firm $i$ before period $t$, by $\tilde{F}^t_j$ the set of firms that have not made an offer to worker $j$ up to (and including) period $t$ and that he weakly prefers to any firm that has made him an offer till then, and by $\mathcal{X}^t$ the set of agents who have exited by period $t$. Let

$$\mathcal{BS}^f_i\left(u^f_{i\cdot}, h^f_{t,i}\right) \equiv \left\{j | j \in \mathcal{S}^f_i\left(u^f_{i\cdot}\right) \cap \tilde{W}^t_i\right\} \setminus \mathcal{X}^t,$$

$$\mathcal{BS}^w_j\left(u^w_{\cdot j}, h^w_{t,j}\right) \equiv \left\{i | i \in \mathcal{S}^w_j\left(u^w_{\cdot j}\right) \cap \tilde{F}^t_j\right\} \setminus \left(\mathcal{X}^t \cup M^F_t\right).$$

**Assumption 2** *Suppose all agents follow decentralized reduced DA strategies.*[32] *Let* $U$ *be in the support of* $G$ *and* $l$ *be an agent of type* $\alpha \in \{f, w\}$. *Assume that in period* $t$, *if* $\alpha = w$, *the worker* $l$ *has not received offers from all but one of the remaining firms. For each* $j, k \in \mathcal{BS}^\alpha_l\left(u^\alpha_l(.), h^\alpha_{t,l}\right)$, *if* $u^\alpha_l(j) > u^\alpha_l(k)$ *and* $k \in S^\alpha_l\left(u^\alpha_l(.), h^\alpha_{t,l}\right)$, *then* $j \in S^\alpha_l\left(u^\alpha_l(.), h^\alpha_{t,l}\right)$.

Assumption 2 simply poses that when all market participants follow decentralized reduced DA strategies, the ordering of offers and matches does not convey information in and of itself to either workers or firms. Put differently, it ensures firms cannot cross out their favorite available plausible stable match partner from the set of perceived plausible stable matches $\mathcal{S}^f_i\left(u^f_{i\cdot}, h^f_{t,i}\right)$ when using updating based on worker matches, exits, and

[32]Any decentralized reduced DA strategy profile leads to the same learning pattern pertaining to stable matches, and so the assumption's requirement is not affected by which particular profile is used.

rejections (generating $\mathcal{BS}_i^f(u_{i\cdot}^f, h_{t,i}^f)$). An analogous intuition holds for workers. An exception occurs when a worker receives $\tilde{F}-1$ offers, where $\tilde{F}$ is the number of firms remaining. Suppose the worker does not receive an offer from his most preferred firm. Then, since the worker is aware that a top-top match must be reached in each period, it must be that the remaining firm has made a top-top match and will exit the market this period. Thus, the worker believes the most-preferred firm of the $\tilde{F}-1$ that had made him an offer is his stable match partner.

As we show in Proposition 4* below, the combination of Assumptions 1 and 2 generalize our notion of richness in the body of the paper, and allows for a wider collection of economies in which decentralized reduced DA strategies are incentive compatible. We therefore introduce:

**Generalized Richness** An economy satisfies *generalized richness* if it satisfies Assumptions 1 and 2.

As in our definition of rich economies in the body of the paper, generalized richness refers to the *support* of potential match utilities. It does not rule out probabilistic updating on the likelihood of different agents being one's stable match in the realized market. While generalized richness is still somewhat restrictive, it is an assumption that is satisfied in several leading examples:

**Examples of Economies that satisfy Generalized Richness**

1. **Full-support economies.** As we prove in Proposition 4 below, a rich economy satisfies generalized richness.

2. **Assortative economies.** Economies in which the distribution of markets is supported on all preference profiles in which, say, firms (similarly workers) have identical ordinal rankings of workers (similarly, firms) are generalized rich. In addition, economies in which the support coincides with all preference profiles in which firms agree on ordinal rankings of workers *and* workers agree on ordinal rankings of firms are generalized rich. For brevity, we omit the proof, which follows the lines of our arguments below.

These examples highlight the idea that generalized richness implies that there is either a strong linkage between agents' supported preferences, the extreme case corresponding to complete information, in which no learning at all takes place during the decentralized market game, or a very weak linkage, the extreme case being a full support economy, so that learning occurs only by eliminating agents who have exited the market or been involved in a rejection.

**Proposition 4*** (Aligned Preferences–Existence)**.** *Suppose the economy $\mathcal{E}$ satisfies generalized richness. For sufficiently high $\delta < 1$, there exists a Bayesian Nash equilibrium in weakly undominated strategies of the decentralized market game that implements the unique stable matching in each supported market.*

**Proof of Proposition** 4* We first show that, for sufficiently high $\delta$, a minimal reduced DA strategy is a best response for a worker whenever 1. all other workers use minimal reduced DA strategies, specifying only plausible stable match partners as acceptable, and 2. firms use mixed decentralized reduced DA strategies.

Indeed, at each period $t$, for sufficiently high $\delta$, a worker cannot benefit by exiting the market whenever a plausible stable partner is still available, nor from accepting an offer from a firm who is not his most preferred plausible stable match. Last, a worker with offers at hand cannot benefit by rejecting a set of firms different than the set of all firms other than his most preferred. This last point follows from the no-cycle property. Indeed, rejection of firms cannot generate the arrival of an offer from a preferred firm, and reduced DA strategies assure that rejected firms will not make future (repeat) offers. In particular, while these three types of deviations could speed up the final match, they can never improve it. With positive probability, the worker will have rejected their best plausible match and lose positive utility as a result, which is never optimal for sufficiently high $\delta$.

We now show that generalized richness ensures that a firm's best responses are within the class of reduced DA strategies whenever 1. workers use a minimal decentralized reduced DA strategy and 2. firms use mixed decentralized reduced DA strategies.

Consider first a firm $i$ that in period $t$ has no outstanding offers, and whose updated strategies suggest worker $j$ as the most preferred stable match. There are two kinds of deviations from a decentralized reduced DA prescription: make no offer, or make an offer to some other worker $k$ who is ranked below $j$. The benefits of such deviations can be either through speeding up the time at which the firm's offer is accepted, or through generating a preferred ultimate match.

If firm $i$ does not make an offer at period $t$, there are three potential implications. First, if making an offer according to any decentralized reduced DA strategy would not have affected market participants' history following period $t$,[33] then the only effect of this deviation could be the prolonging of its match creation. If not making an offer affects certain participants' histories, then due to Assumption 2, this cannot affect the firm's final match. Again, such a deviation can only prolong the timing of that final match. Finally, suppose that the firm's most preferred plausible stable partner receives offer from all other firms remaining on the market at period $t$. Then, the worker will accept the best offer of those immediately, even if he prefers firm $i$ (see our discussion following Assumption 2), in which case firm $i$ is made strictly worse off by the deviation.

Suppose firm $i$ makes an offer to a worker $k$ who is ranked lower than her most preferred plausible stable match $j$. By Assumption 1, the firm could be immediately rejected, in which case she does not benefit. Alternatively, with positive probability, her offer will be held or accepted by $k$ in a market in which she would have otherwise gotten a preferable worker. Such a deviation can never lead to a preferable ultimate match from the incentive compatibility inherent in the firm-proposing DA algorithm. Indeed, note that such a deviation would be tantamount to submitting an untruthful preference list when the firm-proposing DA algorithm is used (as, from Assumption 2, such an offer will not make other participants change their effective rank orderings). However, revealing preferences truthfully is a dominant strategy for firms.

Consider now the restricted centralized market game in which workers' strategy set is confined to minimal decentralized reduced DA strategies, and firms' strategy set is confined to decentralized reduced DA strategies (mixed or pure). Since there is a finite number of firms' decentralized reduced DA strategies, an equilibrium exists in this restricted game, possibly involving mixed strategies. From the above, for sufficiently high $\delta$, the

[33]For instance, in the case in which any such strategy suggests an offer to $j$, who gets matched to another firm in that period with probability 1.

corresponding strategy profile is also an equilibrium in our original decentralized market game, as required. ■

**Proof of Proposition 4**

We show that rich economies satisfy generalized richness. Then, by Proposition 4*, the result follows.

Alignment implies that agents' match utilities are a monotonic transformation of the elements of their respective row or column of the ordinal potential matrix. Without loss of generality, we assume match utilities take values in $\{\varepsilon, 1, 2, \ldots, n\}$, where $n = W$ or $n = F$ depending on whether the agent is a firm or a worker. $\varepsilon \in (0,1)$ corresponds to the utility from remaining unmatched. An equivalent construction follows for any other set of distinct values that contains at least as many elements. Our construction also follows in much the same way if utility values differ across market participants. For any ordinal potential $\Phi$, the mapping from entries $\Phi_{ij}$ to utility values for a firm $i$ is as follows. Let $j_1 = argmax_j \ \Phi_{ij}, j_2 = argmax_{j \neq j_1} \ \Phi_{ij}$, etc. Then, $u^f_{ij_1} = n, u^f_{ij_2} = n-1$, and so on. Worker utility values are mapped similarly.

In what follows, we construct ordinal potentials that rationalize each history an agent observes and guarantee generalized richness. Our arguments are constructive: we identify a set of ordinal potentials consistent with preferences and histories any agent might observe when called upon to act. However, the precise values we place on entries of these ordinal potentials are of no importance—they simply allow us to identify the profile of plausible preference rankings that can then be translated directly to match utilities. Notably, our construction is designed for on-path play. It requires all agents to have followed reduced DA strategies.

First, we consider an arbitrary firm. We begin by establishing restrictions on the ordinal potential $\Phi$ that correspond to the public history, namely the sequence of market exits. It is useful to note that, due to the top-top match property, at least one firm and worker pair exit each period until only firms or only workers remain.

For any firm $i$ and worker $j$ that match and exit in period 1, set $\Phi_{ij} = 0$. Next, for any firm $i$ and worker $j$ that match in period 2, choose an arbitrary firm $i'$, who exited in period 1. Set $\Phi_{i'j} = -j$ and $\Phi_{ij} = -(n+1)$. This restriction rationalizes behavior in these two periods. Indeed, with such an ordinal potential, reduced DA instructs firm $i$ to make an offer to worker $j$ in period 1. Worker $j$ holds the offer but does not accept, in the hope that firm $i'$ will make him an offer. Upon observing firm $i'$'s exit in period 1, worker $j$ accepts his second-best firm $i$. This specification leads to an additional constraint: the remaining elements in the row corresponding to firm $i$ and the column corresponding to worker $j$ must be lower than $-(n+1)$. We continue this construction iteratively through time. In period $k$, for any firm $i$ and worker $j$ pair that exits, we set $\Phi_{ij} = -(k-1)(n+1)$ and $\Phi_{i'j} = -(k-2)(n+1) - j$, where firm $i'$ is an arbitrary firm that exited in the previous period. This specification rationalizes the public information available in the economy at any period.

Consider now firm $i$ at some period $t$, with history $h^f_{t,i}$. The firm knows her own preference ranking, as well as whether workers she has made offers to held or rejected her offers over time. We impose further restrictions on the ordinal potential to accommodate this additional information and rationalize $h^f_{t,i}$. First, the firm's preference ranking implies that the ordering of $\Phi_{ix}$ over $x$ is determined by firm $i$'s preferences.

In the second case, we only consider situations where every worker that firm $i$ has made an offer to has rejected the firm. This is without loss of generality, as otherwise the firm has no strategic choice and the requirements for generalized richness are trivially satisfied.

If worker $j$ held firm $i$'s offer for some number of periods, then exited with a firm $i'$ in period $t' \leq t$, $\Phi_{i'j}$ is specified by the above construction. Namely, $\Phi_{i'j} = -(t'-1)(n+1)$. This implies that firm $i$ would have been chosen by worker $j$ had firm $i'$ been unavailable. Furthermore, it ensures any remaining workers rank below worker $j$ in firm $i$'s preferences. This specification determines $\Phi_{i'j'}$ for any worker $j'$ that exited before $t'$ so that firm $i'$ does not make an early offer to worker $j$. Importantly, the number of rounds that the firm's offer is held is rationalized by the above construction: worker $j$ accepts the offer from firm $i'$ only when firm $i'$ is worker $j$'s most preferred remaining firm.

Similarly, a worker $j$ may hold firm $i$'s offer for some number of periods, then reject the firm's offer without exiting at period $t' \leq t$. With reduced DA strategies, this implies worker $j$ initially preferred firm $i$'s offer to all other received offers. However, in period $t'$, some firm $i'$ made the worker an offer preferable to firm $i$'s offer. Worker $j$ remains in the market only if there exists another firm $i''$ he prefers to firm $i'$ that is a plausible stable match. One path consistent with firm $i$'s observed history has firm $i''$ exiting in period $t-1$. From our specification thus far, $\Phi_{i''j'}$ is determined, where worker $j'$ is the worker with whom firm $i''$ exited. Firm $i$ can also consistently expect worker $j$ to exit with firm $i'$ in period $t$. We thus let $\Phi_{i'j} = -(t-1)(n+1)$ and $\Phi_{ij} = -(t-1)(n+1) - j$. Again, the fact that worker $j$ remains in the market for $t - t'$ periods is accounted for by our specifications up to now.

The restrictions imposed until now are consistent with firm $i$'s sequence of offers through period $t$. We add a restriction: for all workers $j$ to whom firm $i$ has not made an offer, $\Phi_{ij} \leq -t(n+1)$. This constraint is consistent with the requirements placed so far.

For other entries of the potential, we require that remaining workers have not received offers from their most-preferred remaining firms. Additionally, if there is more than one remaining firm, we set the potential so that no remaining workers received offers from all but one firm. Specifically, for each worker $j$ who has not exited the market by period $t$, take a firm $i'$ who exited in period $t-1$ and set $\Phi_{i'j} = -(t-2)(n+1) - j$. Then, each unmatched worker waited for an offer from firm $i'$ in period $t-1$.

Since the economy is rich, at least one market occurring with positive probability has an ordinal potential that satisfies all the restrictions imposed above.

We now show that the restricted set of ordinal potentials that satisfy the above restrictions are sufficient for generalized richness.

For Assumption 1, take any worker $j$ who is not preferred to firm $i$'s most preferred worker and who has not rejected her yet at period $t$. Consider ordinal potentials under which worker $j$ might accept an offer made by firm $i$. Without loss of generality, let firm $i$ prefer worker 1 to worker 2 and so on. Then, indeed, we can add the restrictions $\Phi_{ij} = -n(n+1) - j$ and $\Phi_{i'j} = -n(n+1) - j - i'$ for each firm $i'$ still in the market. With this specification, every remaining worker $j$ immediately accepts offers made by firm $i$, thereby satisfying assumption 1. Richness implies that potentials satisfying these additional restrictions are in the support of firm $i$'s beliefs.

As for Assumption 2, consider an ordinal potential where firm $i$'s most preferred remaining worker $j$ accepts the firm's offer. Set $\Phi_{ij} = -(t-1)(n+1)$. By design, every firm-worker pair with $\Phi_{ij} > -(t-1)(n+1)$ contains at least one agent that has exited in a previous period. Therefore, firm $i$ and worker $j$ constitute a top-top match in period $t$. In particular, worker $j$ will immediately accept firm $i$'s offer. Again, richness implies that potentials satisfying these additional restrictions are in the support of firm $i$'s beliefs.

Consider now a worker $j$ at any arbitrary period $t$. The worker knows his own preference ranking and the sequence of firms that have made him offers up to time $t$, captured by the history $h^w_{t,j}$. As for firms, we find a subset of ordinal potentials that rationalize the worker's preferences and observed history. We will use the following claim.

**Claim** *Suppose all agents follow reduced DA strategies. Assume that, in period $t$, worker $j$ rejects an offer from firm $i$. Then, if in period $t' > t$ firm $i$ exits and worker $j$ remains, there exists a firm $i'$ that has not made an offer to worker $j$ that also exits in period $t'$.*

**Proof of Claim:** Assume $\Phi$ is the governing potential in the realized market. Under reduced DA strategies, if a worker $j$ rejects firm $i$, he must have received an offer from a preferred firm. Suppose that, in period $t'$, the worker holds an offer from firm $i^*$. Let $\tilde{F}$ and $\tilde{W}$ be the sets of firms and workers that have not exited before period $t'$. Any pair in $\arg\max_{(\tilde{\imath},\tilde{\jmath})\in\tilde{F}\times\tilde{W}} \Phi_{\tilde{\imath}\tilde{\jmath}}$ constitutes a top-top match that exits in period $t'$. In particular, there is a firm $i'$ that exists with a worker $j'$. Suppose firm $i'$ made an offer to worker $j$ at some period. From the definition of DA strategies, it follows that $\Phi_{i^*j} > \Phi_{i'j}$ and $\Phi_{i'j} > \Phi_{i'j'}$, in contradiction to firm $i'$ and worker $j'$ being a top-top match in period $t'$.

We restrict ordinal potentials to rationalize exited firm-worker pairs. Namely, we construct an "optimistic" ordinal potential consistent with the public history, where any worker believes his top firm remaining in the market is still achievable. Since all ordinal potentials occur with positive probability, the worker prefers to follow a decentralized reduced DA strategy to avoid risking losing his top firm. Specifically, suppose firm $i$ and worker $j'$ exit in period $t$. We consider two cases. If firm $i$ had made an offer to worker $j$ before, we set $\Phi_{ij'} = -(n+1)^2$. If firm $i$ had not made an offer to worker $j$ before, we set $\Phi_{ij'} = -(t-1)(n+1)$.

In each period, either no firm makes an offer to worker $j$, or at least one firm does. Since the worker has not exited prior to period $t$, he could not have received an offer from his most preferred remaining firm in any prior period.

Suppose available firms $i_1, i_2, \ldots, i_k$ have not made offers to worker $j$ in any period $t' \leq t$ for $t > 1$. Then, worker $j$ infers that each of these firms preferred other available workers in previous periods. Optimistically, set their preference rankings as follows. In period 1, take a matched firm-worker pair, with worker $j_1$. For every firm $i \in \{i_1, i_2, \ldots, i_k\}$ set $\Phi_{ij_1} = -i$. That is, worker $j$ assumes that every such firm made an offer to the exiting worker. In period 2, there must be a worker $j_2$ who exits with a firm that never made worker $j$ an offer. Indeed, if worker $j$ rejected an offer in period 1, such a worker $j_2$ is guaranteed by the claim above. If worker $j$ did not receive any offers in period 1 or is holding his period 1 offer, such a worker $j_2$ is guaranteed from the top-top match property of aligned markets. Then, for every firm $i \in \{i_1, i_2, \ldots, i_k\}$ set $\Phi_{ij_2} = -(n+1) - i$. Repeat this construction up to period $t$, in each period $l < t$ take a worker $j^l$ who exited in period $l$, and set $\Phi_{ij^l} = -(l-1)(n+1) - i$ for all firms $i \in i_1, i_2, \ldots, i_k$. This construction also generates the additional restriction that all remaining potential entries involving two agents who have not yet exited be below $-(t-1)(n+1)$. Note that firms previously rejected by worker $j$ are placed at the "bottom" of the ordinal potential in accordance with their match utilities to ensure that the resulting ordinal potential is consistent with worker $j$'s preferences as described in the previous section.

Suppose worker $j$ has offers from firms $\tilde{\imath}_1, \ldots, \tilde{\imath}_m$ in period $t$, which may include any offer held from period $t$. Without loss of generality, suppose worker $j$ prefers firm $\tilde{\imath}_1$ to firm $\tilde{\imath}_2$ to firm $\tilde{\imath}_3$ and so on. We consider two cases. First, suppose firm $\tilde{\imath}_1$ is worker $j$'s most preferred remaining firm. In that case, reduced DA directs worker $j$ to accept the offer. Since generalized richness does not impose any restrictions in this case, any specification of a consistent potential would do. Second, suppose there exists an available firm $i$ that worker $j$ prefers to firm $\tilde{\imath}_1$. Decentralized reduced DA directs worker $j$ to hold $\tilde{\imath}_1$'s offer and reject the rest. For each rejected firm $i_v = \tilde{\imath}_2, \ldots, \tilde{\imath}_k$, set $\Phi_{i_v j} = -(n+1)^2 + u^w_{i_v j}$. As for firm $\tilde{\imath}_1$ and the most preferred firm $i$, set $\Phi_{\tilde{\imath}_1 j} = -t(n+1)$ and $\Phi_{ij} = -(t-1)(n+1) - i - 1$.

Next, we show that both of the generalized richness assumptions hold for the workers. Assumption 1 holds immediately, as it only places restrictions on firms' beliefs.

Consider Assumption 2. If there exists only one remaining firm, assumption 2 is satisfied automatically. Otherwise, if worker $j$ has received an offer from his most preferred firm, there is no remaining further preferred firm so assumption 2 holds, as suggested above. Last, suppose worker $j$ has not received offers from all but one of the remaining firms and has not received an offer from his most preferred firm. Then, there exist two distinct firms, say $i$ and $i'$, with $i$ preferable to $i'$, which have not made worker $j$ an offer yet. Consistent with the construction above, we can set $\Phi_{ij} = -(t-1)(n+1) - i$ and $\Phi_{i'j} = -(n+1)^2 + u^w_{i'j}$.

The above construction is consistent with the worker's observed history, which implies that $i'$ exits this period, and $i$ makes an offer to worker $j$ next period.

Since for both types of agents Assumptions 1 and 2 are satisfied under reduced DA strategies, richness implies generalized richness and the proposition's claim follows. ■

## References


Agarwal, N. (2015): "An empirical model of the medical match," *American Economic Review*, 105, 1939–78.

Agranov, M., A. Dianat, L. Samuelson, and L. Yariv (2021): "An Experimental Study of Incomplete Information in Matching Markets," *mimeo*.

Akbarpour, M., S. Li, and S. O. Gharan (2020): "Thickness and information in dynamic matching markets," *Journal of Political Economy*, 128, 783–815.

Alcalde, J., D. Pérez-Castrillo, and A. Romero-Medina (1998): "Hiring procedures to implement stable allocations," *Journal of Economic Theory*.

Alcalde, J. and A. Romero-Medina (2000): "Simple mechanisms to implement the core of college admissions problems," *Games and Economic Behavior*, 31, 294–302.

Arnosti, N., R. Johari, and Y. Kanoria (2014): "Managing congestion in decentralized matching markets," in *Proceedings of the fifteenth ACM conference on Economics and computation*, 451–451.

Ashlagi, I., Y. Kanoria, and J. D. Leshno (2017): "Unbalanced random matching markets: The stark effect of competition," *Journal of Political Economy*, 125, 69–98.

Baccara, M., S. Lee, and L. Yariv (2020): "Optimal dynamic matching," *Theoretical Economics*, 15, 1221–1278.

Baccara, M. and L. Yariv (2022): "Dynamic Matching," *Online and Matching-Based Market Design*.

Bikhchandani, S. (2017): "Stability with one-sided incomplete information," *Journal of Economic Theory*, 168, 372–399.

Blum, Y., A. E. Roth, and U. G. Rothblum (1997): "Vacancy chains and equilibration in senior-level labor markets," *Journal of Economic theory*, 76, 362–411.

Burdett, K. and M. G. Coles (1997): "Marriage and class," *The Quarterly Journal of Economics*, 112, 141–168.

Chade, H., J. Eeckhout, and L. Smith (2017): "Sorting through search and matching models in economics," *Journal of Economic Literature*, 55, 493–544.

Clark, S. (2006): "The uniqueness of stable matchings," *Contributions in Theoretical Economics*, 6, 1–28.

Dai, X. and M. Jordan (2021): "Learning in Multi-Stage Decentralized Matching Markets," *Advances in Neural Information Processing Systems*, 34.

Diamantoudi, E., E. Miyagawa, and L. Xue (2015): "Decentralized matching: The role of commitment," *Games and Economic Behavior*, 92, 1–17.

Dur, U. M., P. A. Pathak, F. Song, and T. Sönmez (2018): "Deduction Dilemmas: The Taiwan Assignment Mechanism," Tech. rep., National Bureau of Economic Research.

Echenique, F., R. Gonzalez, A. Wilson, and L. Yariv (2022): "Top of the Batch: Interviews and the Match," *American Economic Review: Insights*, forthcoming.

Echenique, F., A. Robinson-Cortes, and L. Yariv (2012): "An experimental study of decentralized matching," *mimeo*.

Eeckhout, J. (1999): "Bilateral search and vertical heterogeneity," *International Economic Review*, 40, 869–887.

——— (2000): "On the uniqueness of stable marriage matchings," *Economics Letters*, 69, 1–8.

Ehlers, L. and J. Massó (2007): "Incomplete information and singleton cores in matching markets," *Journal of Economic Theory*, 136, 587–600.

Fernandez, M. A., K. Rudov, and L. Yariv (2022): "Centralized Matching with Incomplete Information," *American Economic Review: Insights*, forthcoming.

Fréchette, G. R., A. E. Roth, and M. U. Ünver (2007): "Unraveling yields inefficient matchings: evidence from post-season college football bowls," *The RAND Journal of Economics*, 38, 967–982.

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

Grenet, J., Y. He, and D. Kübler (2021): "Decentralizing centralized matching markets: Implications from early offers in university admissions," *mimeo*.

Gutin, G. Z., P. R. Neary, and A. Yeo (2021): "Unique Stable Matchings," *mimeo*.

Haeringer, G. and M. Wooders (2011): "Decentralized job matching," *International Journal of Game Theory*, 40, 1–28.

Hall, R. E. and A. B. Krueger (2012): "Evidence on the incidence of wage posting, wage bargaining, and on-the-job search," *American Economic Journal: Macroeconomics*, 4, 56–67.

Immorlica, N. and M. Mahdian (2005): "Marriage, honesty, and stability," in *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, Society for Industrial and Applied Mathematics, 53–62.

Kojima, F. and P. A. Pathak (2009): "Incentives and stability in large two-sided matching markets," *American Economic Review*, 99, 608–27.

Konishi, H. and M. Sapozhnikov (2008): "Decentralized matching markets with endogenous salaries," *Games and Economic Behavior*, 64, 193–218.

Liu, Q. (2020): "Stability and Bayesian Consistency in Two-Sided Markets," *American Economic Review*, 110, 2625–2666.

Liu, Q., G. J. Mailath, A. Postlewaite, and L. Samuelson (2014): "Stable matching with incomplete information," *Econometrica*, 82, 541–587.

Monderer, D. and L. S. Shapley (1996): "Potential games," *Games and economic behavior*, 14, 124–143.

Niederle, M. and A. E. Roth (2003): "Unraveling reduces mobility in a labor market: Gastroenterology with and without a centralized match," *Journal of political Economy*, 111, 1342–1352.

——— (2009): "Market culture: How rules governing exploding offers affect market performance," *American Economic Journal: Microeconomics*, 1, 199–219.

Pais, J. (2008): "Incentives in decentralized random matching markets," *Games and Economic Behavior*, 64, 632–649.

Pais, J., A. Pintér, and R. F. Veszteg (2012): "Decentralized matching markets: a laboratory experiment," *Unpublished paper*.

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

——— (1989): "Two-sided matching with incomplete information about others' preferences," *Games and Economic Behavior*, 1, 191–209.

——— (1991): "A natural experiment in the organization of entry-level labor markets: regional markets for new physicians and surgeons in the United Kingdom," *The American economic review*, 415–440.

——— (2008): "What have we learned from market design?" *Innovations: Technology, Governance, Globalization*, 3, 119–147.

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, 748–780.

Roth, A. E. and M. Sotomayor (1992): "Two-sided matching," *Handbook of game theory with economic applications*, 1, 485–541.

Roth, A. E. and J. H. Vande Vate (1990): "Random paths to stability in two-sided matching," *Econometrica*, 1475–1480.

Shimer, R. and L. Smith (2000): "Assortative matching and search," *Econometrica*, 68, 343–369.

Sørensen, M. (2007): "How smart is smart money? A two-sided matching model of venture capital," *The Journal of Finance*, 62, 2725–2762.

Suh, S.-C. and Q. Wen (2008): "Subgame perfect implementation of stable matchings in marriage problems," *Social Choice and Welfare*, 31, 163–174.

Voorneveld, M. and H. Norde (1997): "A characterization of ordinal potential games," *Games and Economic Behavior*, 19, 235–242.