NBER WORKING PAPER SERIES

DECENTRALIZED MATCHING WITH ALIGNED PREFERENCES

Muriel Niederle
Leeat Yariv

Working Paper 14840
http://www.nber.org/papers/w14840

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2009

We are grateful to Attila Ambrus, Federico Echenique, Guillaume Haeringer, Fuhito Kojima, and Al Roth for many useful conversations and suggestions. This work is supported by NSF and the Sloan Foundation. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Decentralized Matching with Aligned Preferences
Muriel Niederle and Leeat Yariv
NBER Working Paper No. 14840
April 2009
JEL No. C78

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 match. In the presence of market frictions and preference uncertainty, harsher assumptions on the richness of the economy have to be made in order for decentralized markets to generate stable outcomes in equilibrium.

Muriel Niederle
Department of Economics
579 Serra Mall
Stanford University
Stanford, CA 94305-6072
and NBER
niederle@stanford.edu

Leeat Yariv
Division of Humanities & Social Sciences, Caltech
1200 E. California Blvd
Pasadena, CA 91125
lyariv@hss.caltech.edu

1. Introduction

1.1 Overview

The theoretical literature on two sided matching markets has focused predominantly on the analysis of outcomes generated in centralized markets. There are many examples in which two sided matching markets are centralized (e.g., the medical residency match, school allocations, the U.S. market for reform rabbis, etc.). Nonetheless, many markets are not fully centralized (for instance, college admissions in the U.S., the market for law clerks, junior economists, and so on). Furthermore, almost all centralized markets are preceded by decentralized opportunities for participants to match. Understanding the outcomes generated by decentralized markets is therefore important to the design of institutions, both fully decentralized ones, as well as ones followed by centralized procedures.[1] The current paper offers a first step in that direction.

The key feature of centralized clearinghouses that empirically predicts their continued use in a market is whether they produce a stable outcome.[2] From a theoretical perspective, the set of stable outcomes coincides with the core in environments such as the ones we study. The assumption that markets achieve the core is, in fact, utilized in empirical work that uses stability constraints to deduce market participants' characteristics.[3] Our results provide conditions under which a (non-cooperative) decentralized market game yields a core outcome, and the econometric identifying assumption of stability is likely to hold. When the structure of the economy does not guarantee existence of a stable outcome, there may be room for market design if stability is taken as a goal.[4]

We study a simple model of a decentralized *market game* in which firms make directed

[1] See Roth (1984, 2008). For the recent literature on differences in outcomes between centralized clearinghouses and decentralized markets see Frechette, Roth, and Unver (2007) and Niederle and Roth (2003). They 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 gastroentorology fellows. On rare occasions, decentralized bargaining prior to the centralized match is prohibited by design, such as in some residency matches in the UK (Roth, 1991).

[2] A *stable match* is a pairing of workers and firms (where some workers and some firms may be left alone), 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. For empirical evidence on the importance of stability for a centralized clearinghouse see Roth (1991).

[3] See, e.g., Hitsch, Hortacsu, and Ariely (2006), Sorensen (2007) and Lee (2009).

[4] This includes the introduction of a centralized clearinghouse (Roth, 1984) or efforts to ease information transmission in the presence of significant frictions and congestion (Coles and Niederle, 2009).

offers to workers. In our setup, 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 frictions in the economy.

In more detail, we focus on markets in which firms can employ up to one worker, who can work for at most one firm. We assume that a match with any agent is preferred to remaining unmatched. We consider environments in which there is a unique stable match. This allows us to sidestep coordination problems. Furthermore, we concentrate on a special class of preferences that guarantee uniqueness of the stable match that we term *aligned preferences.* Aligned preferences require that the preferences of firms and workers can be represented by a joint ordinal potential. Alignment implies that there is always a firm-worker pair that are each other's first choice. Many prominent cases studied in the literature entail aligned preferences. For instance, alignment is guaranteed whenever firms and workers generate revenue they split in fixed proportions, or when all participants on one side of the market share the same preference ranking of the other side's participants (that can have arbitrary preferences).

In our decentralized market game, in every period, each firm can make up to one offer to a worker of her choice if she does not already have an offer held by a worker in the market. Workers can accept, reject, or hold on to an offer, in which case it is available also in the next period. Firms and workers share a common discount factor, and receive their match utilities as soon as they are matched (by having an offer accepted), or leave the market. This allows us to study the effect of frictions through discounting.

We study two cases of information agents have about their own and each others' preferences: (a) *complete information* in which all market participants are fully informed of the realized match utilities; (b) *private information* in which each agent is fully informed only of her own match utilities. While the extant literature has mostly focused on complete information environments (see below), we find the case of private information particularly important from an empirical point of view.

Throughout our analysis, we concentrate on equilibria in weakly undominated strategies.

Under complete information, all agents can compute the stable match and we show that the stable match is an equilibrium outcome in the market game we analyze. Still, there may be equilibria that yield unstable matches, highlighting a first contrast with findings from the case in which agents use a centralized market, where all equilibria in weakly undominated strategies yield the stable match. Nonetheless, simple refinements (namely, iterated elimination of weakly dominated strategies) restore uniqueness of the stable match as an equilibrium outcome.[5]

When information is private, in a frictionless economy, stable matches may 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: firms make offers to workers in order of their preferences, and workers accept offers when they are made by their most preferred firms. The fact that aligned preferences assure there is always a firm-worker pair that are each others' first choice guarantees that at least one firm and worker are matched in every period. In particular, such a strategy profile leads to a market match in finite time.

However, as soon as there are frictions in the market, agents may have incentives to deviate from these strategies to speed up the matching process, or affect market participants' learning regarding their expected stable matches. Our analysis then follows two steps. First, ignoring incentive compatibility constraints, we characterize the class of strategy profiles that generate stable outcomes. Second, we identify economies under which at least one of these (possibly mixed) profiles is incentive compatible. The message that emerges is that when the economy is sufficiently rich in terms of possible market realizations, there is a Bayesian Nash equilibrium that implements the unique stable match.

1.2 Related Literature

There are several recent theoretical advances that inspect market outcomes as consequences of a dynamic interaction. Haeringer and Wooders (2009) and Pais (2008) consider the case of complete information, and restrict firms' strategies in that they cannot make offers to workers who had rejected them previously. Haeringer and Wooders (2009) study a game similar to ours in which firms can only make exploding offers (that have to be accepted or rejected right away). In Pais (2008) one firm is chosen randomly each period to make an offer. She characterizes the set of (ordinal) subgame perfect equilibria and shows that outcome multiplicity may arise

[5] This is in line with laboratory observations of decentralized markets with complete information, in which stability frequently achieved, see Echenique, Katz, and Yariv (2009) and Niederle and Roth (2009).

even when the underlying market has a unique stable match.[6,7]

Our setup bears some similarity to that considered in the search literature on matching (e.g., Burdett and Coles, 1997, Eeckhout, 1999, and Shimer and Smith, 2000). There, each period, workers and firms randomly encounter each other, observe the resulting match utilities, and decide jointly whether to pursue the match and leave the market or to separate and wait for future periods. As in our setting, equilibrium outcomes depend on the distribution of match utilities. Unlike our setting, the perceived distribution of potential partners does not change with time, and each side of the market solves an option value problem.

Regarding our assumptions, alignment is reminiscent of some identified sufficient conditions for uniqueness of a stable match (Clark, 2006 and Eeckhout, 2000).

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 finite sets of firms and workers, respectively, and $U = \left\{\left\{u_{ij}^f\right\}, \{u_{ij}^w\}\right\}$ 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 no match. 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: (1) firms and workers have strict preferences. That is, for any firm i, $u_{ij}^f \neq u_{ij'}^f$ for any $j, j' \in \mathcal{W} \cup \varnothing$ and for any worker j, $u_{ij}^w \neq u_{i'j}^w$ for any $i, i' \in \mathcal{F} \cup \varnothing$;

[6]The role of commitment in dynamic games as ours with complete information is highlighted in Diamantoudi, Miyagawa, and Xue (2007). Similarly, Blum, Roth and Rothblum (1997) study dynamics when the firms' but not the workers' commitments are binding.

[7]There is also some work analyzing endogenous salaries in decentralized markets with complete information and limited dynamics see Konishi and Sapozhnikov (2008). For the analysis of wages in a centralized clearinghouse, see Crawford and Knoer (1981) and Kelso and Crawford (1982). Related empirical work (such as Choo and Siow, 2006 and Fox, 2008) has used constraints derived from stability with transferable utility to estimate underlying preference parameters. 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 matches (see also Alcalde, Pérez-Castrillo, and Romero-Medina, 1998).

[8]Cardinal utilities are required to trade off matchings at different points in time and examine the impacts of discounting.

(2) match utilities are strictly positive – for all $i \in \mathcal{F}$ and $j \in \mathcal{W} \cup \varnothing$, $u^f_{ij} > 0$, and similarly for all $j \in \mathcal{W}$ and $i \in \mathcal{F} \cup \varnothing$, $u^w_{ij} > 0$; (3) all agents prefer to be matched over remaining unmatched. Therefore, for any $i \in \mathcal{F}$, all workers $j \in \mathcal{W}$ are *acceptable*, $u^f_{ij} > u^f_{i\varnothing}$, and for any $j \in \mathcal{W}$, all firms $i \in \mathcal{F}$ are *acceptable*, $u^w_{ij} > u^w_{\varnothing j}$.

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 utility levels $U \in \mathcal{U}$.

A *match* 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 μ. A *blocking pair* for a match μ 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 match is *stable* if it is not blocked by any pair (note that matches are not blocked by individuals, as all agents prefer to be matched over remaining unmatched).

Gale and Shapley (1962) showed that any market has a stable match, and provided an algorithm that identifies one. In the *firm proposing deferred acceptance 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, firms whose offer got rejected in the last step make an offer to the most preferred worker who has not rejected them 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, any firm either has their offer held by a worker, or has been rejected by all workers. Once the algorithm ends, held offers turn into matches.

The resulting match is the firm optimal stable match, i.e., for any firm it is the stable match that is not dominated by any other stable match. It is in turn the least preferred stable match for workers. Similarly, there always exists a worker optimal stable match, which is the least preferred stable match for firms. In general, these two matches can be different, and many other stable matches can exist.

In this paper we do not want problems of coordination on a specific stable match to be the hurdle to the existence of an equilibrium yielding a stable outcome in the market game (the details of which we soon describe). Therefore, we only consider markets $\mathcal{M} = (\mathcal{F}, \mathcal{W}, U)$

that have a unique stable match denoted by μ_M. In such markets a centralized mechanism, to which agents submit rank ordered lists of preferences, has an equilibrium that always generates the stable outcome, while this is not the case when there are multiple stable matches (see e.g., Roth and Sotomayor, 1990, and section 3 in the paper).

2.2 Aligned Preferences

The literature has not identified general necessary and sufficient conditions on a market for the stable match to be unique. Throughout our analysis, we will focus on a class of preferences, termed *aligned preferences*, that guarantees uniqueness of the stable match.

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

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

The notion of ordinal potential is analogous to that of a potential in two player games in which agents' match utilities replace the payoff matrix $\left(U^w, U^f\right) = \left((u^w_{ij}, u^f_{ij})\right)_{i\in\mathcal{F},j\in\mathcal{W}}$ (see Monderer and Shapley, 1994).

Conceptually, 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 Φ. Let $\hat{U}^f = \hat{U}^w = (\Phi_{ij})_{i\in\mathcal{F},j\in\mathcal{W}}$, then $\hat{U}^f$ and $\hat{U}^w$ capture the same ordinal preferences over partners as U^f, U^w.

Most applied papers implicitly assume that preferences are aligned (see, e.g., Sorensen, 2007, and references therein). In particular, two prominent examples of aligned preferences are the following:

Examples of Aligned Preferences

1. Firms and workers have a joint production output they share in fixed proportions when they are matched. That is, there exists a number $\alpha > 0$ such that for all $(i, j) \in \mathcal{F} \times \mathcal{W} : u^w_{ij} = \alpha u^f_{ij} > 0$. Here, $\Phi = (\Phi_{ij})_{i\in\mathcal{F},j\in\mathcal{W}}$ defined as $\Phi_{ij} = u^f_{ij}$ for all i, j, serves as an ordinal potential.

2. Suppose one side of the market has the same preference ranking over participants on the other side of the market (the preferences of whom can be arbitrary). To see that these preferences are aligned, assume without loss of generality that it is the firms who share the same preference ordering 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 $\tilde{u}_{ij}^w$, such that for all $(i,k) \in \mathcal{F}, j \in \mathcal{W}$: $0 < \tilde{u}_{ij}^w < 1$, and $\tilde{u}_{ij}^w < \tilde{u}_{kj}^w \Leftrightarrow u_{ij}^w < u_{kj}^w$. Then $\Phi = (\Phi_{ij})_{i\in\mathcal{F},j\in\mathcal{W}}$ with $\Phi_{ij} = j + \tilde{u}_{ij}^w$ for all i,j is an ordinal potential.

As it turns out, Clark (2006) assures uniqueness of the stable match when preferences are aligned:

Remark (Alignment – Uniqueness) *When preferences are aligned, there is a unique stable match.*

Providing the intuition for uniqueness will be substantially easier once we present other properties of aligned preferences.

One important attribute of preference alignment is 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, then 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 j_1 by rejecting i_2 can trigger an offer from a preferred i_1 :

$$u_{i_1j_1}^w > u_{i_2j_1}^w, \quad u_{i_2j_1}^f > u_{i_2j_2}^f, \quad u_{i_2j_2}^w > u_{i_3j_2}^w, \quad u_{i_3j_2}^f > u_{i_3j_3}^f, ..., \quad u_{i_nj_n}^w > u_{i_1j_n}^w, \quad u_{i_1j_n}^f > u_{i_1j_1}^f$$

Note that such a chain would be equivalent to having a cycle in the payoff matrix $\left(U^w, U^f\right)$. We then say that preferences satisfy the **no cycle property.** As it turns out, the characterization of potential games by Voorneveld and Norde (1997) assures that preference alignment is equivalent to the no cycle property.

A market $(\tilde{\mathcal{F}}, \tilde{\mathcal{W}}, \tilde{U})$ is a sub-market of $(\mathcal{F}, \mathcal{W}, U)$ if $\tilde{\mathcal{F}} \subseteq \mathcal{F}$, $\tilde{\mathcal{W}} \subseteq \mathcal{W}$, and $\forall i,j \in (\tilde{\mathcal{F}} \cup \varnothing) \times (\tilde{\mathcal{W}} \cup \varnothing) \setminus \{\varnothing, \varnothing\}$, $\tilde{u}_{ij}^w = u_{ij}^w$ and $\tilde{u}_{ij}^f = u_{ij}^f$. A second implication of preference alignment is that for any sub-market $(\tilde{\mathcal{F}}, \tilde{\mathcal{W}}, \tilde{U})$ 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 $\tilde{\mathcal{F}}$. We say that preferences satisfy the **top-top match** property.[9] Intuitively, when preferences are aligned, the original market, as well as any sub-market, has an ordinal potential. Suppose $\tilde{\Phi}$ is an ordinal potential of the sub-market $(\tilde{\mathcal{F}}, \tilde{\mathcal{W}}, \tilde{U})$. 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 $(\tilde{\mathcal{F}}, \tilde{\mathcal{W}}, \tilde{U})$. Hence preference alignment implies the top-top match property.

This observation provides the intuition for the fact that a stable match is unique. In fact, it can be identified through a recursive process. In the initial step, find the firm-worker pairs that constitute $top - top$ matches, pairs (i,j) at which the ordinal potential $\Phi = (\Phi_{ij})$ achieves a local maximum. The corresponding firms and workers must be matched to each other for the match to be stable. The remaining firms and workers form a market with aligned preferences and we can continue recursively. By construction, this procedure generates the unique stable match.

Finally, preferences admit a **stable blocking pair** whenever for any match $\mu \neq \mu_M$ there exists a blocking pair (i,j) such that $\mu_M(i) = j$. Note that when preferences are not aligned (even if there is a unique stable match μ_M), any unstable match μ will allow for a blocking pair $(i,j) \in \mathcal{F} \times \mathcal{W}$ such that $u^f_{i\mu_M(i)} \geq u^f_{i\mu(i)}$ and $u^w_{\mu_M(j)j} \geq u^w_{\mu(j)j}$. That is, both i and j prefer the outcome in the stable match μ_M to the outcome in μ (see Roth and Sotomayor, 1990). When preferences are aligned there exists a stable blocking pair: a blocking pair (i,j) for which $\mu_M(i) = j$. Indeed, suppose $\mu \neq \mu_M$. Going through the recursive process described above (to illustrate uniqueness), at some stage a discrepancy must arise between μ_M and μ. At that stage, a match that occurs under μ_M does not get formed. In fact, the corresponding worker and firm form a stable blocking pair. Proposition 1 summarizes all these claims.[10]

Proposition 1 (Alignment – Properties)

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

2. *If preferences are aligned then the top-top match and stable blocking pair properties hold.*

[9] The top-top match property is referred to as α-reducibility in Clark (2006).

[10] Alignment is a stronger property than the property proposed by Eeckhout (2000) in Theorem 1. This is easy to see, as Eeckhout's condition does not imply the top-top match property.

Note that there are markets that fulfill the top-top match property, and the stable blocking pair property, but in which preferences are not aligned (and, therefore, do not exhibit the no cycle property). Consider for example the following market with three firms and three workers and match utilities:

$$(u^f_{ij}, u^w_{ij})_{i\in\mathcal{F}, j\in\mathcal{W}} = \begin{array}{|c|c|c|}\hline 3,5 & 1,1 & 5,2 \\ \hline 4,4 & 2,3 & 7,7 \\ \hline 1,1 & 7,2 & 6,3 \\ \hline \end{array}.$$

It easy to check that the top-top match and stable blocking pair properties hold. Nonetheless, U exhibits a rejection chain, and hence does not correspond to aligned preferences: $u^w_{11} > u^w_{21}$, $u^f_{21} > u^f_{22}$, $u^w_{22} > u^w_{32}$, $u^f_{32} > u^f_{33}$, $u^w_{33} > u^w_{13}$, and $u^f_{13} > u^f_{11}$.

2.2 A Decentralized Market

For a given economy $\{(\mathcal{F}, \mathcal{W}, U)\}_{U\in\mathcal{U}}$ together with a distribution G over utility realizations U, we analyze the following *market game*. The economy, together with the distribution G, is common knowledge to all agents. At the outset of the game, the market is realized according to the distribution G. Firms make offers over time, indexed by $t = 1, 2, ...$ and workers react to them. 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 only the offers they received themselves. Each worker j who has received an offer from firm i can accept, reject, or hold the offer. Once an offer is accepted, worker j is matched to firm i. Workers can also decide to 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 got matched.

Eligible firms are firms i 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.

We consider market games without frictions, and market games with frictions, which take the form of discounting. If a firm i is matched to worker j at time t, firm i receives $\delta^t u^f_{ij}$ and worker j receives $\delta^t u^w_{ij}$, 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 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 δ 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).

The fact that time is valuable as described through discounting provides a major obstacle to the decentralized market game reaching a stable outcome. In the conclusions we address other ways in which congestion and market frictions might arise, e.g., fixed costs for making offers, and other possible rules of the market game.

We focus on two configurations of information in the economy. The simplest is that of *complete information* in which both firms' and workers' match utilities are common knowledge.[11] This is the case that most of the literature on decentralized matching has tackled. Note that in this case, both firms and workers can deduce the stable outcome μ_M.

The second information structure we analyze is that of *private information.* In that setup, each agent is informed only of her own match utilities. In the case of private information, for an equilibrium of the market game to be a stable outcome, information has to be transmitted to allow agents to deduce their stable match partner.

Each of these information structures defines a Bayesian game where type spaces correspond to the available private information of each agent. The equilibrium notion we use is that of Bayesian Nash equilibrium. When information is complete, type spaces are singletons and we essentially characterize the Nash equilibria of the corresponding game.

Our analysis concentrates on equilibria of decentralized market games in which all agents use weakly undominated strategies. For both information structures, weak undominance 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, by definition, offers always lead to a higher payoff).

[11] The crucial assumption in the analysis of the complete information case is that each agent knows all other agents' preference *ordering*.

2. When $\delta < 1$, a worker who receives an offer from his most preferred unmatched firm accepts it immediately and, similarly, if only agents on one side of the market are unmatched, they exit immediately.

Note that restriction 2 is due to the fact that all payoffs are strictly positive.

3. CENTRALIZED MATCHING

Before analyzing decentralized markets, we analyze the case of a centralized clearinghouse (such as ones used by medical markets and many school districts) to which agents simultaneously submit preferences. The centralized clearinghouse uses an algorithm that produces a stable outcome given the submitted preferences. We show that, in our environment, with a centralized clearinghouse it is always possible to elicit preferences of agents in a way that yields the stable match as an equilibrium outcome. To ease analogies between centralized and decentralized markets, we assume that agents in a centralized market report match utilities that are then translated into ordinal preferences. That is, each agent submits a vector of positive match utilities.[12] Technically, this is equivalent to having agents simply report ordinal preferences directly.

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.[13] 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 $j \neq \varnothing$, $u_{ij}^f = u_{i\varnothing}^f$ implies $jP(u_{i\cdot}^f)i$.

We define a *deferred acceptance mechanism* as a mechanism in which each agent l of type $\alpha \in \{f, w\}$ reports their match utilities u_l^α simultaneously (after receiving all private information). The mechanism then computes the corresponding ordinal preferences P and associates them with the stable match induced by the firm proposing deferred acceptance algorithm on P. The payoffs of firms and workers correspond to their match utilities given by their match partner. Throughout the paper, we will use the shorthand of DA for the label of "deferred acceptance".

[12] The restriction to positive numbers is made only for presentation simplicity. The sufficient restriction is that the set of available reports contains as many elements as the maximal number of different match utilities.

[13] Note that even though agents never experience indifference in their realized match utilities, they may still report indifferences.

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 match corresponds to the unique stable match in each market of the economy (see Roth and Sotomayor, 1990). That is,

Lemma (Centralized Matching)

1. *For any economy with a unique market all Nash equilibria in weakly undominated strategies of the game associated with the DA mechanism yield the unique stable match* μ_M.

2. *For any economy there exists a Bayesian Nash equilibrium in weakly undominated strategies of the game associated with the DA mechanism such that the corresponding match is the unique stable match* μ_M *in each market.*

It is important to note that even though implementing the stable match is always possible through the centralized clearinghouse, the stable match is not necessarily the unique equilibrium outcome in the presence of uncertainty (see Niederle and Yariv, 2009).

4. Complete Information

We start by analyzing economies in which all participants are informed of the realized market. That is, there is complete information regarding all match utilities, and all agents can compute the stable match. In particular, achieving the stable match can be done in one period. Intuitively, consider strategies where each firm makes an offer to its stable match partner, or exits the market if it is unmatched under the unique stable match μ_M. Each worker accepts his best available offer in period 1, and if he receives no offers, exits the market. This profile constitutes an equilibrium in weakly undominated strategies that yields the match μ_M.

Ruling out weakly dominated strategies is not sufficient to guarantee uniqueness. First, there may still be multiple equilibria generating μ_M. Indeed, for sufficiently high discount factors, an alternative way involves emulating the DA algorithm. Since this profile may entail several periods of market activity, it would generate different equilibrium payoffs.

Furthermore, there may be outcomes generated by equilibria in weakly undominated strategies that do not coincide with the stable match μ_M as the following example illustrates.

Example 1 (Multiplicity with Complete Information). Consider 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 match: $\mu_M(Fi) = Wi$ for all i. Consider the following profile of weakly undominated strategies yielding the match μ: $\mu(F1) = W2, \mu(F2) = W1, \mu(F3) = W3$ and $\mu(F4) = W4$ (corresponding to the underlined entries in the matrix). 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 offer. Off equilibrium, 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 (as long as they're not 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 in case all workers left the market. Furthermore, in period 2, any worker who does not receive offers exits immediately, 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.

The crux of the problem generating the multiplicity above is that elimination of weakly dominated strategies provides little restraint on off-equilibrium behavior. Specifically, $F3$ could "punish" $F2$ for not making an offer to $W1 = \mu(F2)$ in period 1.[14] Iterated elimination

[14] 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.

of weakly dominated strategies rules out such strategies, and in fact, guarantees that the stable match is the unique equilibrium outcome, as the following proposition summarizes.

Proposition 2 (Complete Information) *For any economy, there exists a Nash equilibrium in weakly undominated strategies that yields the stable match for each realized market. Furthermore, the stable match of each realized market 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 sub-market. Since the corresponding workers realize the top-top matches in the original market are formed in period 1, iterated elimination of weakly dominated strategies assures that they accept their top-top matches in the remaining market, and 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 the unique stable match of the market being implemented in one period.[15]

Note that this construction hinges on the fact that all agents are completely informed of the realized market, and hence each firm and worker can compute the stable match. Proposition 2 shows that a robust non-cooperative market game equilibrium results necessarily in the unique stable match (the unique core outcome in this market). In what follows we illustrate the impact of preference uncertainty in the economy.

5. Economies with Uncertainty

For a decentralized market to reach a stable outcome, sufficient information has to be transmitted to ensure that (i) workers only accept offers from firms that are their stable match and (ii) firms make offers to those workers. Furthermore, the decentralized market has to allow for this information to be transmitted in an incentive compatible way.

There are three channels through which information flows in the 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

[15] The uniqueness of equilibrium outcomes is due to preference alignment. In fact, Niederle and Yariv (2009) show that without preference alignment, equilibria surviving iterated elimination of weakly dominated strategies may generate multiple outcomes, even when information is complete.

to those offers (unless offers are accepted, in which case that information becomes public). The third component of information is time – all participants are aware of the period they are in.

Recall that alignment implies the $top - top$ match property. Hence, if firms and workers follow strategies that resemble the DA algorithm (namely, firms make offers to workers in order of their preferences and workers hold on to their best available offer and only accept an offer only from their most preferred available firm), then in every period some agents are matched and public information is transmitted. Furthermore, this process yields the stable outcome.

The main hurdle for establishing stability is that these DA strategies may not be incentive compatible. First, agents may have an incentive to speed up the timing of matches. Second, market activity throughout the game allows participants to *learn* about the realized market. Since updating affects the final matching outcomes, agents may have incentives to deviate from DA strategies in order to manipulate the updating process of other participants.

5.1 Learning in a Decentralized Market

Before investigating strategies of firms and workers, we describe the information agents have at each 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 the market by themselves), 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 that 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 (where we use the notational convention that an offer to no worker is denoted as an offer to $\varnothing$ that is immediately rejected), and the (timed) set of agents that have left the market:

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

In addition, at each period t, suppose workers $j_1, ..., j_{k(t-1)}$ rejected offers from firm i in periods

$1, ..., t-1$. Denote by $\tilde{W}_i^t = \left\{j | j \notin \left\{j_1, ..., j_{k(t-1)}\right\}\right\}$ the set of workers that have not rejected firm i yet.

Each unmatched worker acts in the interim stage of each period t and observes a history that consists of all (timed) offers he received, a (timed) sequence of offers he held, the (timed) set of agents that have left the market prior to t, and the set of firms that left the market at time t:[16]

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

In addition, at each period t, suppose firms $i_1, ..., i_{k(t)}$ made offers to worker j in periods $1, ..., t$. Denote by $\tilde{F}_j^t = \left\{i | u_{ij}^w \geq \max\left\{u_{lj}^w | l = i_1, ...i_{k(t)}\right\}\right\}$ the set of firms that have not made an offer to worker j yet and that he weakly prefers to any firm that has made him an offer thus far.

Agents use the observed history to update the posterior on the realized markets, and the resulting potential stable match partners. For a given prior distribution G over possible utility levels $U \in \mathcal{U}$, for any private information $u_l^\alpha(.)$ of agent l of type $\alpha \in \{f, w\}$ regarding the realized market, let $G(u_l^\alpha(.))$ denote the posterior distribution over utility realizations. Let $\mathcal{S}_l^\alpha(u_l^\alpha(.)) = \{\mu(U)(l) \mid U \in \text{supp } G(u_l^\alpha(.))\}$ denote the set of all ex ante potential stable match partners of agent l. That is, agents that could conceivably be part of a stable match, under the distribution over market match utilities updated by the private information $u_l^\alpha(.)$. Analogously, given the strategies played by all agents, $\mathcal{S}_l^\alpha(u_l^\alpha(.), h_{t,l}^\alpha)$ denotes the set of all interim potential stable match partners given agent's l available information at t.

5.2 Frictionless Economies

Suppose there is no discounting, i.e., $\delta = 1$. Then, one way in which information may be transmitted in the market is if agents simply follow **DA strategies**. 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 only once the offer is from the firm that yields the highest match utility given the set of firms that are still unmatched. These prescriptions are followed by all agents after any detectable deviations as well.

[16] An offer of firm i to worker j that is held from period t to t' is recorded as an offer made in periods $t, t+1, ..., t$' that is held by the worker in each of these periods. We use a similar convention for workers.

Proposition 3 (No Discounting) *Suppose $\delta = 1$, then DA strategies constitute a Bayesian Nash equilibrium in weakly undominated strategies and yield the stable match.*

Intuitively, when all 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. In particular, the process stops in finite time. When $\delta = 1$, the timing of matches is of no importance to market participants and unilateral deviations cannot generate a better match.

5.3 Economies with Frictions

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 $u_{ij}^f = u_{ij}^w$ for all i, j. Match utilities are given as follows:

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

Firm 2 knows that worker 2 is the 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 lose a period. Firms may therefore be tempted not to make all offers in order of their preferences, but rather concentrate on offers to potential stable match partners. Similarly, workers may accept an offer from their highest potential stable match partner, even if more preferred firms are still unmatched.

Before we attack the general problem of sufficient incentive compatible information transmission in decentralized markets, we start by analyzing some minimal conditions strategies have to satisfy in a centralized mechanism in order to generate the stable match. This allows us to temporarily avoid incentive compatibility hurdles that are due to interim learning. That is, our first step is a generalization of the Lemma in Section 3, which illustrated the existence of a Bayesian Nash equilibrium yielding the stable match (namely, one involving agents submitting their preference profile truthfully). The second step of our analysis entails a characterization of economies in which emulating the class of identified strategies (from step 1) in a decentralized market is incentive compatible for all participants.

Certainly, if there is any hope of achieving the stable match (in a centralized setting) for any market realization, agents must declare potential stable matches acceptable. Furthermore, consider, e.g., the firms. Since the centralized mechanism achieves the firm optimal stable match for the submitted preferences, changing the ranking of agents that are preferred to the stable match would not change the resulting match in the centralized market. However, it is crucial that any agents ranked above *any* potential stable match are, in fact, preferred to that stable match. These restrictions suggest the following class of strategies, termed **reduced DA strategies**.

Formally, for each agent $l \in \mathcal{F} \cup \mathcal{W}$ who submits utilities v corresponding to preferences $P(v)$, let $P(v, \varnothing) = \{k : kP(v)\varnothing\} \cup \{\varnothing\}$ where $P(v, \varnothing) \subseteq \mathcal{W} \cup \{\varnothing\}$ if $l \in \mathcal{F}$ and $P(v, \varnothing) \subseteq \mathcal{F} \cup \{\varnothing\}$ if $l \in \mathcal{W}$. That is, $P(v, \varnothing)$ is the union of no partner (i.e., unmatched) and the set of match partners l strictly prefers to being unmatched given $P(v)$. Denote $u_l^\alpha(k) = u_{lk}^f$ if $\alpha = f$ (and $l \in \mathcal{F}$) and $u_l^\alpha(k) = u_{kl}^w$ otherwise.

Definition (Reduced DA Strategies) *Let $\mathcal{E}$ be an economy and assume agents participate in a centralized match: each agent $l \in \mathcal{F} \cup \mathcal{W}$ submits utilities v corresponding to preferences $P(v)$. Agents use* reduced DA strategies *if for each $l \in \mathcal{F} \cup \mathcal{W}$ with $\alpha \in \{f, w\}$:*

1. $S_l^\alpha\left(u_l^\alpha(.)\right) \subseteq P(v, l)$
2. *For each $k, m \in S_l^\alpha\left(u_l^\alpha(.)\right)$, $kP(v)m \Leftrightarrow u_l^\alpha(k) > u_l^\alpha(m)$.*
3. *For each $k \in P(v, l) \backslash S_l^\alpha\left(u_l^\alpha(.)\right)$ and each $m \in S_l^\alpha\left(u_l^\alpha(.)\right)$, $kP(v)m \Rightarrow u_l^\alpha(k) > u_l^\alpha(m)$.*

The intuitive interpretation of the three conditions required by reduced DA strategies is the following. First, any potential stable match (using the agent's private information on match utilities) is declared acceptable. Second, potential stable matches are ranked truthfully. Third, rankings of agents who are not potential stable matches above potential stable matches must be truthful with respect to the stable matches.

We now show that reduced DA strategies in a sense place minimal requirements for securing stability in centralized markets. Note that the conditions of reduced DA strategies depend only on the *support* of stable matches, and not their precise likelihood of occurrence. We consider economies with a finite number of potential markets, and so strategy profiles that generate the stable match 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 used strategy is a reduced DA strategy.[17]

Proposition 4 (Centralized Aligned Economies with Discounting)

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 match.*

2. *Suppose there is an agent $a \in \mathcal{F} \cup \mathcal{W}$ who 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.*

Proposition 4 illustrates the effectiveness of reduced DA rules in generating the stable match outcome. Part 2 of the Proposition highlights the necessity of the conditions imposed by reduced DA strategies for implementing stable outcomes in any economy.

When moving from a centralized mechanism to a decentralized market, we need to translate reduced DA strategies to strategies in decentralized markets. For firms the translation is straightforward. Whenever firm i would submit a reduced DA strategy v, firm i does the following in the decentralized market. 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). When firm i gets rejected by the last

[17] 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 $S_l^\alpha(u_l^\alpha(.))$ is identical when M is realized in either economy. Then, if agent l uses a reduced DA rule, they must use the *same* reduced DA strategy in both $\mathcal{E}_1$ and $\mathcal{E}_2$, whenever observing $u_l^\alpha(.)$.

acceptable worker, the last worker who is still unmatched and has a higher utility than firm i itself 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, and 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, he has to accept the best available offer. 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 the offer from his most preferred unmatched firm. The translation of a reduced DA strategy v to a decentralized market will be captured by "threshold firms". At each point in time, a worker accepts the best available offer whenever he receives an offer that he likes at least as much as his most preferred unmatched firm 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. Beyond that, 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.

A second element to account for when moving from a centralized to a decentralized market is that agents may use strategies in which actions depend on the history of play. Specifically, note that reduced DA strategies depend on the set of potential stable match partners. In decentralized markets, given the strategy profile of all other agents, each firm or worker can recalculate and possibly refine their set of potential 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 them to take information they have accumulated into account. For a firm this implies that the firm makes an offer to workers that are at least as good as her most preferred unmatched potential stable match who has not rejected her yet. For a worker this implies that the best offer that is weakly preferred to his least preferred potential stable match is not rejected. Furthermore, upon receiving an offer at least as good as his most preferred unmatched firm according to v, the worker immediately accepts that offer. The worker rejects all offers that are not as good as the least preferred potential stable match, and only holds one offer. Note that for any agent $l \in \mathcal{F} \cup \mathcal{W}$ of type $\alpha \in \{f, w\}$, given the actions of all other agents, the actions of agent l using a decentralized reduced DA strategy up to any period t can be described by a single reduced DA strategy using $\mathcal{S}_l^\alpha \left(u_l^\alpha(.), h_{t,l}^\alpha\right)$ as the set of potential stable match partners.[18]

In centralized markets, we have seen that reduced DA strategies impose minimal restrictions on strategies to achieve stability and are incentive compatible. However, the decentralized counterpart of many such strategies is generally not incentive compatible in markets with frictions, as agents care about *when* matches are created. This may create incentives to make less offers than those required by decentralized reduced DA strategies. Furthermore, the fact that agents update the set of potential stable match partners using the history of play opens room for a different class of manipulations intended to affect the learning that occurs in the market (and, consequently, agents' perception of their conceivable stable match partners). Proposition 4 above suggested that decentralized reduced DA strategies are natural candidates for strategy profiles that yield stable outcomes. Our goal now is to characterize the class of economies for which a subset of these strategies are incentive compatible.

When all agents in the market follow decentralized reduced DA strategies, firms may at times be able to speed up the process by altering the ranking of agents. Intuitively, suppose that all other players use strategies that implement the stable outcome. There are economies in which a firm's offer to a worker j who is not its first choice worker will be accepted only if that worker is actually its stable match partner. Then the firm may have an incentive to make that offer first, in order to speed up the timing of its match, as such an "out of order" offer entails no risk of "wrong acceptance" if all other agents use strategies that implement the stable outcome. Below is a simple example of such a case in which no equilibrium implements the stable outcome with certainty.[19]

Example 2 (Timing of Matches). Consider an economy with two firms $\{F1, F2\}$ and two workers $\{W1, W2\}$ with $u_{ij}^f = u_{ij}^w$ for each of 6 potential markets, described by the

[18] As time progresses, learning can only lead to the elimination of previously perceived potential stable matches. Thus, the conditions imposed by reduced DA strategies are weakened in the interim stages.

[19] Note that workers cannot speed things up by holding on to offers that are not from potential stable match partners. However, holding such offers is not necessarily strictly harmful to the worker, since the no cycle property assures that the rejection of an offer cannot trigger a chain yielding a more preferred offer.

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 will show that there are no equilibria in weakly undominated strategies that always implement the stable outcome.

U_3 and U_4 guarantee that $F1$ sometimes makes an offer to $W1$ in any equilibrium when $W1'$s match utilities are $(3,4)$.[20] Similarly, U_5 and U_6 guarantee that $W2$ with match utilities $(6,5)$ will in equilibrium sometimes receive no offer in period 1, but only in period 2.[21] From now on, we focus on U_1 and U_2.[22]

Note that $W1$ 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 given $F1$'s match utilities $(3,6)$, 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, $F2$, that is aware U_2 is realized, makes an offer to $W1$ in period 1, who will accept 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. Note that along the equilibrium path $F1$ makes an offer to $W1$ with match utilities of $(3,4)$ only in U_3, when $F1$ is the stable match, hence $W1$ 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 to $W2$ 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 δ). The idea is that $F1$ can assure that

[20] 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 $F1$). Note that in U_3 it must be the case that $F2$ makes an offer to $W2$ that gets immediately accepted. Therefore, in period 2, if $W1$ rejected the offer of $F1$, $F1$ can remake that offer, in which case $W1$ has to acept it whenever using a weakly undominated strategy. Hence, in equilibrium, $F1$ best responds by making an offer to $W1$ whenever her match utilities are $(3,2)$.

[21] 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.

[22] In particular, $F1$ observes $(3,6)$, $W1$ observes $(3,4)$, and they cannot distinguish between U_1 and U_2.

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. First, note that this implies that $W1$ has to accept the offer from $F1$ with positive probability, $m \in (0,1]$.[23] In order for the market to always yield a stable outcome, it has to be the case that $F2$ makes an offer to $W1$ with probability 1 in period 1 when U_2 prevails, which implies that $W1$ has to accept the offer from $F1$ whenever he receives that offer. 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 5δ, which is bigger than 4 for $\delta > 4/5$. Hence, for large enough δ, $F1$ 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 $1 - p$.

Assumption 1 rules out economies as in the example above. It makes sure 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 potential stable match 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. Either the firm will have her offer held or accepted, as it is better than the stable match of the worker in the realized market.[24] Alternatively, in case the worker uses a decentralized reduced DA strategy that specifies firms that are less preferred than all potential stable matches as unacceptable, the firm will be rejected immediately. Then, making an offer to that worker triggers no chain of offers, rejections, or acceptances, and as such has no benefit over not making an offer at all.

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

[24] Therefore, in that realized market, the worker is a worse match for the firm than her stable match, as otherwise the worker and firm would constitute a blocking pair.

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

Decentralized reduced DA strategies require that in each period t and for each history $h^f_{t,i}$ firm i makes an offer that corresponds to some reduced DA strategy given information $(u^f_{i\cdot}, h^f_{t,i})$. This implies that firm i makes an offer to a worker who is at least as good for her as her highest ranked unmatched potential stable match partner who has not rejected her yet. If workers always hold the best available offer that is at least as desirable as their lowest potential stable match partner (unless they accept an offer), then the set of potential stable matches after some history is given by:

$$\begin{aligned}
\mathcal{S}^f_i\left(u^f_{i\cdot}, h^f_{t,i}\right) &\subseteq \mathcal{BS}^f_i\left(u^f_{i\cdot}, h^f_{t,i}\right) \equiv \left\{j | j \in \mathcal{S}^f_i\left(u^f_{i\cdot}\right) \cap \tilde{W}^t_i\right\} \setminus \mathcal{X}^t, \\
\mathcal{S}^w_j\left(u^w_{\cdot j}, h^w_{t,j}\right) &\subseteq \mathcal{BS}^w_j\left(u^w_{\cdot j}, h^w_{t,j}\right) \equiv \left\{i | i \in \mathcal{S}^w_j\left(u^w_{\cdot j}\right) \cap \tilde{F}^t_i\right\} \setminus \left(\mathcal{X}^t \cup M^F_t\right).
\end{aligned}$$

The inclusion above may certainly be strict. For instance, consider an economy comprised only of U_1 and U_2 of Example 2. In that case, if firms use a decentralized reduced DA strategy, worker $W1$ who receives an offer from $F1$ only in period 1 can infer that he will not receive another offer and that $F1$ is his stable match partner. In particular, he could accept that offer, even though it is not from his first choice firm. In fact, such examples hinge on there being effectively only two relevant firms and workers. However, such examples can be embedded in larger markets.

Definition (Top Sub-economy) *Suppose all agents use decentralized reduced DA strategies. A top sub-economy for worker j_1 consists of two firms i_1 and i_2 and an additional worker j_2 such that in period t, worker j_1 has sufficient information to infer that:*

1. *for any $j \in \{j_1, j_2\}$: $i_1, i_2 \in \mathcal{S}^w_j\left(u^w_{\cdot j}, h^w_{t,j}\right)$ and no other potential stable match is ranked higher than i_1* or *i_2.*

2. *Both* i_1 and i_2 *either have an outstanding offer to* $j \in \{j_1, j_2\}$ *or make an offer to them in period* t.

The generalization of the example above imposes a restriction on updating when a top sub-economy is reached. Indeed, suppose all agents use decentralized reduced DA strategies and that in period t worker j has a top sub-economy with firms $\{i_1, i_2\}$. If i is the most preferred firm among those with an offer to j in period t then $S_j^w\left(u_{\cdot j}^w, h_{t,j}^w\right) = \{i\}$ (even when i is only the second choice firm among $\{i_1, i_2\}$).

As it turns out, in general cases in which $\mathcal{S}_l^\alpha(u_l^\alpha(.), h_{t,l}^\alpha) \subsetneq \mathcal{BS}_l^\alpha(u_l^\alpha(.), h_{t,l}^\alpha)$ for an agent l of type $\alpha \in \{f, w\}$, decentralized reduced DA strategies may be sensitive to manipulations. Specifically, timing of offers can be used to affect the set of perceived potential stable partners. The following example illustrates the power of such manipulations in an economy in which there does not exist a Bayesian Nash equilibrium in weakly undominated strategies that always yields the stable outcome.

Example 3 (Manipulability of Offer Timing). 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:[25]

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

Suppose there exists an equilibrium in weakly undominated strategies that always implements the stable match.

U_3 guarantees that in such an equilibrium $F3$ with match utilities $(6, 8, 2)$ always makes an offer to $W2$ in period 1.[26] Furthermore, U_4 guarantees that in any such equilibrium $F1$ with match utilities $(5, 4, 3)$ makes an offer to $W1$ with probability 1.[27] From now on we

[25]Note that the economy satisfies Assumption 1.

[26]Such an offer will be immediately accepted in U_3. Furthermore, if $W2$ makes an offer to $W1$ with some probability $p > 0$ in period 1, then in U_3, $W1$ is aware that the stable match partner is $F1$ yelding match utility of 5, so $W1$ will accept that offer, yielding an unstable outcome.

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

concentrate on 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). So 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 $\mathbf{U}_2$ with probability $\mathbf{q} \in (\mathbf{0}, \mathbf{1}]$. When $F1$ makes an offer to $W3$, $F1$ receives a payoff of 4, as $W3$ accepts 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 potential stable match partners at $t = 1$ is $\{F1, F3\}$ when either is realized. The crux of the problem is the fact that $W1$ cannot be certain whether he will receive his best offer in period 1 or period 2. $W1$ tries to infer that from offers received in period 1. However, offers can then be manipulated. The 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. In what follows we will identify economies that are consistent with these restrictions.

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, with the caveat that there is no top sub-economy. Formally,

Assumption 2 *Suppose all agents follow decentralized reduced DA strategies. 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 no top sub-economy. For each $j, k \in \mathcal{BS}_l^\alpha \left(u_l^\alpha(.), h_{t,l}^\alpha\right)$, if $u_l^\alpha(j) > u_l^\alpha(k)$ and $k \in S_l^\alpha \left(u_l^\alpha(.), h_{t,l}^\alpha\right)$, then $j \in S_l^\alpha \left(u_l^\alpha(.), h_{t,l}^\alpha\right)$.*

Assumption 2 assures firms cannot cross out their favorite available potential stable match from the set of perceived potential stable matches $\mathcal{S}_i^f(u_{i\cdot}^f, h_{t,i}^f)$ when using updating based on worker matches, exits, and rejections (generating $\mathcal{BS}_i^f(u_{i\cdot}^f, h_{t,i}^f)$). Similarly for workers. An exception occurs when a worker is in a top sub-economy. In that case, when a worker receives only one offer, he believes the offering firm is his stable match.[28]

Assumptions 1 and 2 prove crucial in assuring incentive compatibility of decentralized reduced DA strategies. We therefore introduce:

Definition (Rich Economy) *An economy is* rich *if it satisfies Assumptions 1 and 2.*

Note that 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 richness is certainly restrictive (we shall return to some important examples later on), it is an assumption that is satisfied in several leading examples:

Examples of Rich Economies

1. **Complete Information Economies.** Under complete information, agents know their stable match at the outset (it is a singleton).

2. **Full Support Economies.** For a fixed set of firms $\mathcal{F}$ and workers $\mathcal{W}$, let $\mathcal{U}(\Pi)$ denote the set of all aligned match utilities in which each match utility is taken from a set of potential payoffs Π. Consider the economy in which full support is put on elements of $\mathcal{U}(\Pi)$. When Π contains enough elements, say, $2WF + W + F$ (so that some elements of $\mathcal{U}(\Pi)$ are such that all match utilities in the market are different from one another), the economy is rich.

The examples highlight the idea that richness essentially implies that there is either a lot of correlation between agents' realized preferences (the extreme case corresponding to complete information), in which no learning at all takes place during the decentralized market game, or very little correlation (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.

[28] Any decentralized reduced DA strategy profile leads to the same learning pattern regarding stable matches, and so the main requirement of Assumption 2 is not affected by which particular profile is used. Which reduced DA strategy profile is used may, however, affect the time at which a top sub-economy occurs.

Before stating our existence result, we introduce a **mixed decentralized reduced DA strategy** which puts positive probability only on pure decentralized reduced DA strategies. Recall that decentralized reduced DA strategies are weakly undominated. The equilibrium implementability of stable outcomes is captured by the following proposition.

Proposition 5 (Aligned Preferences – Existence) *Suppose the economy is rich. For sufficiently high δ, there exists a Bayesian Nash equilibrium in weakly undominated strategies of the decentralized market game in which workers use decentralized reduced DA strategies and firms use mixed decentralized reduced DA strategies. Furthermore, the outcome is the unique stable match.*

The proof of Proposition 5 follows three steps. First, we set the decentralized reduced DA strategy of workers to be minimal, so that they only rank potential stable match partners. When firms use arbitrary mixed decentralized reduced DA strategies we show that: 1. Workers are best responding, and 2. Each firm's best responses is within the set of decentralized reduced DA strategies (though not necessarily coinciding with those prescribed by the original profile considered). Second, we consider a restricted (finite) market game in which agents' strategies are restricted so that workers use minimal decentralized reduced DA strategies and firms use arbitrary decentralized reduced DA strategies. We find an equilibrium in that game. From the first step, it follows that this is an equilibrium in the full game. These two steps establish the existence of an equilibrium in which all agents use (possibly mixed) decentralized reduced DA strategies. The third step uses Proposition 4 to deduce that in any such equilibrium, the market outcome is stable. Note that these strategies survive iterated elimination of weakly dominated strategies.

Proposition 5 implies that, as long as learning in the market is restricted in terms of what can be deduced from the mere timing of events, stability can arise as an equilibrium outcome.[29]

6. Conclusions and Extensions

We analyzed when a decentralized market game in which firms make offers and workers react to them allows for an equilibrium in weakly undominated strategies that yields the stable match. This is the case when the economy consists of a single market, and hence all agents know who

[29] Uniqueness, however, is generally not guaranteed.

their stable match partner is. It is also the case when there are no market frictions (taking the form of discounting in our model). When there are both uncertainty and discounting, the economy needs to be sufficiently rich and discounting has to be insubstantial enough for there to be an equilibrium that yields a stable outcome.[30]

The paper studied the link between a cooperative concept, the core (which is equivalent to the set of stable outcomes) and (Bayesian) Nash equilibrium in a market game. In our environments there is always a way to implement the stable outcome as a (Bayesian) Nash equilibrium using a centralized clearinghouse to which agents (strategically) report (ordinal) preferences. Hence, the paper emphasizes the differences between markets organized in a decentralized as opposed to a centralized way.

A first difference is that in a decentralized market game, actions, i.e., offers, (temporary) acceptances, and rejections can depend on past histories. The consequences of such contingent actions are easily seen in the case of complete information. Decentralized markets allow for Nash equilibria in which the outcome may be unstable, even though a centralized clearinghouse implements only stable outcomes through Nash equilibrium. Our alignment assumption assures that simple refinements restore uniqueness in decentralized markets. However, when preferences are not aligned, multiplicity of equilibrium outcomes withstands these refinements (see Niederle and Yariv, 2009). Indeed, for general preferences, markets with a unique stable match may contain sub-markets admitting more than one stable match. This may introduce incentives for firms to deviate from the equilibrium that yields the stable outcome of the whole market.

A second difference between a centralized clearinghouse and a decentralized market game concerns whether sufficient information is transmitted to allow agents to infer the stable outcome. In a decentralized market game enough information has to be transmitted in an incentive compatible way so that: first, workers end up with offers from the right firms and, second, workers know when to accept offers. Here too alignment is important. Indeed, consider even the frictionless decentralized market game that always admits an equilibrium yielding the stable outcome. The construction of such an equilibrium hinges on the fact that, due to alignment, in every period there is a top-top match. That is, there is a firm and worker pair

[30] When discounting is substantial, in the presence of uncertainty richness is not sufficient to guarantee an equilibrium that yields the stable outcome.

that exits the market. This makes the market conclude in finite time. When preferences are not aligned, it is less obvious whether workers perceive when the market is over, potentially providing firms with incentives to manipulate these perceptions by withholding offers for a few periods.[31]

Finally, a centralized clearinghouse can implement in equilibrium the stable outcomes in the presence of both uncertainty and frictions (the latter playing no role in the centralized setting). In decentralized markets, stable outcomes can be generated through Bayesian Nash equilibria only when the economy is rich. As it turns out, the sufficiency of our richness assumption relies on preference alignment. Indeed, the stable blocking pair property of aligned preferences assures that firms making offers only to potential stable match partners is sufficient to eventually generate the stable outcome. This is not necessarily the case for general preferences (again, see Niederle and Yariv, 2009).

Our analysis also highlights the importance of frictions, captured by discounting or the probability of a market breakdown. We note that frictions can take a variety of forms that would lead to similar conceptual insights. For instance, frictions could take the form of costly offers. Suppose generating an offer costs an amount of c. All of our existence results would carry through where, in analogy to a vanishing discounting factor, c would become sufficiently small.

We posed a particular normal form game as a model for decentralized markets. Namely, we assumed that in every period firms first simultaneously decide whether and to whom to make an offer. Workers collect all their offers and decide whether to accept, reject, hold offers or exit the market. We stress that these market rules are not only natural on realistic grounds, but also (possibly consequently) shared by the vast majority of papers on decentralized markets. Furthermore, our conceptual results are robust to several modifications. Indeed, the model can be directly extended to allow for markets in which only a fraction of firms are able to make offers in every period. The results can also be directly translated to a symmetric world in which workers make offers and firms collect them. Nonetheless, we view the market game analyzed in this paper as a starting point.

Ultimately, the paper shows that when studying markets, it is generally crucial to under-

[31] Without alignment, this discussion suggests the need for fuller monitoring of actions in the market to generate stable outcomes in equilibrium, even without discounting (see Niederle and Yariv, 2009).

stand market characteristics that go beyond the identification of market participants and their preferences. Indeed, it is important to describe markets in detail, in terms of the information available to participants and the plausibility of frictions, in order to be able to predict which outcomes they may achieve. This, in turn, implies that channels by which information can be transmitted among market participants can be a critical element of market design.

7. Appendix - Proofs

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

Let $M^{(1)} = \{(i,j) \in \mathcal{F} \times \mathcal{W} \text{ s.t. } (i,j) \in \arg\max_{(i,j)\in\mathcal{F}\times\mathcal{W}} \Phi_{ij}\}$. Define $\mathcal{F}_M^{(1)} = \{j|\exists i \in \mathcal{W}$ s.t. $(i,j) \in M^{(1)}\}$ and $\mathcal{W}_M^{(1)} = \{i|\exists j \in \mathcal{F}$ s.t. $(i,j) \in M^{(1)}\}$. Denote $\mathcal{F}^{(2)} = \mathcal{F} \setminus \mathcal{F}_M^{(1)}$, $\mathcal{W}^{(2)} = \mathcal{W} \setminus \mathcal{W}_M^{(1)}$. The sub-market corresponding to $\mathcal{F}^{(2)}$ and $\mathcal{W}^{(2)}$ has aligned preferences and the same ordinal potential and the procedure above can be replicated. That is, for any k, define $M^{(k)} = \{(i,j) \in \mathcal{F}^{(k)} \times \mathcal{W}^{(k)} \text{ s.t. } (i,j) \in \arg\max_{(i,j)\in\mathcal{F}^{(k)}\times\mathcal{W}^{(k)}} \Phi_{ij}\}$, $\mathcal{F}_M^{(k)} = \{j : \exists i \in \mathcal{W}$ s.t. $(i,j) \in M^{(k)}\}$, and $\mathcal{W}_M^{(k)} = \{i : \exists j \in \mathcal{F}$ s.t. $(i,j) \in M^{(k)}\}$, $\mathcal{F}^{(k+1)} = \mathcal{F} \setminus \mathcal{F}_M^{(k)}$, and $\mathcal{W}^{(k+1)} = \mathcal{W} \setminus \mathcal{W}_M^{(k)}$. Note that the unique stable match μ_M can then be identified using recursion: for each k, whenever $(i,j) \in M^{(k)}$, $\mu_M(i) = j$ (and any unmapped remaining agents are unmatched under μ_M).

Proof of Proposition 1.

1. Notice that 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 $\Phi = (\Phi_{ij})$. For any sub-market $(\tilde{\mathcal{F}}, \tilde{\mathcal{W}}, \tilde{U})$, 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 μ is a match different than the unique stable match μ_M. Reconstruct μ_M as above. Consider the smallest k such that there is a pair (i,j) that is matched under μ_M and not under μ. In that case, (i,j) blocks μ and $(i,j) \in M^{(k)}$ form a top-top match in the sub-market corresponding to $\mathcal{F}^{(k)}$, $\mathcal{W}^{(k)}$. As this is the first discrepancy between μ_M and μ 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 μ_M must make offers that get accepted immediately to their corresponding match partner under μ_M. It then follows that firms or workers that are not matched under μ_M must exit the market in period 1. In particular, any profile surviving iterated elimination of weakly dominated strategies must then entail the match μ_M being realized in period 1. ■

Proof of Proposition 3.

Since preferences are aligned, 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 match in finite time. From the convergence of the Gale-Shapley algorithm to a stable match it follows that DA strategies yield the stable match.

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. Therefore, if i a potential stable match partner, then 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. 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 match). The second potential deviation of a worker is the acceptance of an offer that is not from his most preferred unmatched firm. However, with time workers are made better off, 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 match in the market (assuming all other agents follow the DA strategies) is μ. The match μ has the property that the set of firms F', who prefer this match to the stable match μ_M, is not 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 μ to a firm in F' (see, e.g., Roth and Sotomayor, 1990). However, since i^* and j^* follow DA strategies, i^* must have made an offer to j^*, in contradiction. ■

Proof of Proposition 4.

1. Assume all agents use a reduced DA rule and suppose $\mathcal{E}$ is an economy with a market realization in which the outcome μ' is different than the stable match μ_M. By Proposition 1, there exists a pair (i^*, j^*) that blocks μ 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 a themselves) as preferred to a potential stable match 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 potential stable match as unacceptable, then whenever the market in which the unique stable match entails that individual match for a, the centralized outcome is not stable.

Suppose that $a \in \mathcal{W}$ ranks a potential stable match i below a firm i' who is not a potential stable match when observing $u^w_{\cdot a}$ and the set of potential 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 matches 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:[32]

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 match, while under $\tilde{U}$, i' and a are part of the stable match.

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 match is an element $i'' \in S \setminus \{i\}$.

If the stable match 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 match is achieved under $\tilde{U}$, it must be 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 match is established. Indeed, the centralized mechanism generates a stable match 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 potential stable match i below a less preferred potential stable match i' when observing $u^w_{\cdot a}$ and the set of potential stable matches is S. Furthermore, analogous constructions follow when agent a is a firm that does not follow a reduced DA rule.

[32] 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 rubric corresponds to the firm's preference and the second number to the appropriate worker:

$$U: \quad \begin{array}{c|c|c|} & a & j \\ \hline i & u^w_{ia}, u^w_{ia} & u^w_{ia} - 1, u^w_{ia} - 1 \\ \hline i\prime & u^w_{i'a}, u^w_{i'a} & u^w_{i'a} - 1, u^w_{i'a} - 1 \\ \hline \end{array} \qquad \tilde{U}: \quad \begin{array}{c|c|c|} & a & j \\ \hline i & \tilde{u}^f_{i\varnothing} - 1, \tilde{u}^w_{\varnothing a} - 1 & \tilde{u}^f_{i\varnothing} - 2, \tilde{u}^w_{\varnothing j} - 1 \\ \hline i\prime & u^w_{i'a}, \tilde{u}^w_{\varnothing a} + 1 & u^w_{i'a} - 1, \tilde{u}^w_{\varnothing j} + 1 \\ \hline \end{array}.$$

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 σ_j. If σ_j is also a reduced DA strategy, then by part 1 above the outcome is unchanged, in contradiction. Suppose then that σ_j yields a match μ 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 μ such that $\mu_M(i^*) = j^*$. First, it is clear that $j^* \neq j$ since j strictly prefers μ to μ_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^*) are a blocking pair to μ. ■

Proof of Proposition 5. We first note that for sufficiently high δ, whenever all other workers use minimal reduced DA strategies, specifying only potential stable match partners as acceptable, and firms use mixed decentralized reduced DA strategies, using a minimal reduced DA strategy is a best response for a worker.

Indeed, at each period t, for sufficiently high δ, a worker cannot benefit by exiting the market whenever a potential stable match is still available, nor from accepting an offer from a firm who is not his most preferred potential stable match. Last, a worker having offers at hand cannot benefit by rejecting a set of firms different than the set of all firms but his most preferred (in case it is more preferred than his lowest potential stable match). From the no cycle property, 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, such deviations cannot speed up matches, nor alter positively the ultimate match.

We now show that richness assures that whenever workers use a minimal decentralized reduced DA strategy and firms use mixed decentralized reduced DA strategies, a firm's best responses are within the class of 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: (1) make no offer, or (2) 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.

Regarding (1), 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 the match. Finally, suppose that the firm's most preferred potential stable match has a top sub-economy at period t that contains firm i. If the worker receives an offer from the other firm in his top sub-economy, he will accept that offer immediately, even if he prefers firm i, in which case firm i is strictly worse off.

Regarding (2), suppose firm i makes an offer to a worker k who is lower ranked than her most preferred potential stable match j. By Assumption 1, either the firm will be immediately rejected, in which case she does not benefit, or 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 better 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 (possibly mixed) exists in this restricted game. From the above, for sufficiently high δ, the corresponding strategy profile is also an equilibrium in our original decentralized market game, as required. ■

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

8. References

Alcalde, José, David Pérez-Castrillo, and Antonio Romero-Medina. 1998. "Hiring procedures to implement of stable allocations," *Journal of Economic Theory*, **82**, 469-480.

Alcalde, José and Antonio Romero-Medina. 2000. "Simple Mechanisms to Implement the Core of College Admissions Problems," *Games and Economic Behavior*, **31**, 294–302.

Blum, Yosef, Alvin E. Roth, and Uri Rothblum. 1997. "Vacancy chains and equilibrium in senior-level labor markets," *Journal of Economic Theory*, **76**, 362-411.

Burdett, Ken and Melvyn G. Coles. 1997. "Marriage and Class," *The Quarterly Journal of Economics*, **112(1)**, 141-168.

Choo, Eugene and Aloysius Siow. 2006. "Who Marries Whom and Why," *Journal of Political Economy*, **114(1)**, 175-201.

Clark, Simon. 2006. "The Uniqueness of Stable Matchings," *BE Press - Contributions to Theoretical Economics*, **6(1)**, Article 8.

Coles, Peter and Muriel Niederle. 2009. "Signaling in Matching Markets," mimeo.

Crawford, Vincent P., and Elsie Marie Knoer. 1981. "Job Matching with Heterogenous Firms and Workers," *Econometrica,* **49(3)**, 437-450.

Diamantoudi, Effrosyni, Eiichi Miyagawa, and Licun Xue. 2007. "Decentralized Matching: The Role of Commitment," mimeo.

Echenique, Federico, Gabriel Katz, and Leeat Yariv. 2009. "An Experimental Study of Decentralized Matching," mimeo.

Eeckhout, Jan. 1999. "Bilateral Search and Vertical Heterogeneity," *International Economic Review*, **40(4)**, 869-887.

Eeckhout, Jan. 2000. "On the uniqueness of stable marriage matchings," *Economic Letters*, **69**, 1-8.

Frechette, Guillaume, Alvin E. Roth, and M. Utku Unver. 2007. "Unraveling Yields Inefficient Matchings: Evidence from Post-Season College Football Bowls," *RAND Journal of Economics*, **38(4)**, 967-982.

Fox, Jeremy. 2008. "Estimating Matching Games with Transfers," mimeo.

Gale, David, and Lloyd S. Shapley. 1962. "College Admissions and the Stability of

Marriage," *American Mathematical Monthly*, **69**, 9-15.

Haeringer, Guillaume and Myrna Wooders. 2009. "Decentralized Job Matching", mimeo.

Hitsch, Guenter J., Ali Hortacsu, and Dan Ariely. 2006. "What Makes You Click? Mate Preferences and Matching Outcomes in Online Dating," mimeo.

Kelso, Alexander S., and Vincent P. Crawford. 1982. "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, **50(6)**, 1483-1504.

Konishi, Hideo and Margarita Sapozhnikov. 2006. "Decentralized Matching Markets with Endogeneous Salaries," mimeo.

Lee, Soohyung. 2009. "Preferences and Choice Constraints in Marital Sorting: Evidence From Korea," mimeo.

Monderer, Dov and Lloyd S. Shapley. 1994. "Potential Games," *Games and Economic Behavior*, **14**, 124-143.

Niederle, Muriel and Alvin E. Roth. 2003. "Unraveling reduces mobility in a labor market: Gastroenterology with and without a centralized match," *Journal of Political Economy*, **111(6)**, 1342-1352.

Niederle, Muriel, and Alvin E. Roth. 2009. "Market Culture: How Rules Governing Exploding Offers Affect Market Performance", *American Economic Journal: Microeconomics*, forthcoming.

Niederle, Muriel, and Leeat Yariv. 2009. "Matching through Decentralized Markets," mimeo.

Pais, Joana V. 2008. "Incentives in decentralized random matching markets," *Games and Economic Behavior*, **64**, 632-649.

Roth, Alvin 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.

Roth, Alvin E. 1991. "A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K.," *American Economic Review*, **81**, 415-440.

Roth, Alvin E. 2008. "What have we learned from market design ?," *Economic Journal*, **118**, 285-310.

Roth, Alvin E. and Marilda Sotomayor. 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis.* Econometric Society Monograph Series, Cambridge University Press.

Shimer, Robert and Lones Smith. 2000. "Assortative matching and search," *Econometrica*, **68(2)**, 343-369.

Sorensen, Morten. 2007. "How Smart is Smart Money? A Two-Sided Matching Model of Venture Capital," *Journal of Finance*, **62**, 2725-2762

Voorneveld, Mark and Henk Norde. 1997. "A Characterization of Ordinal Potential Games," *Games and Economic Behavior*, **19**, 235-242.