

UCLA

UCLA Previously Published Works

Title

Information Cascades

Permalink

<https://escholarship.org/uc/item/8nm2t5qw>

Authors

Bikhchandani, Sushil
Hirshleifer, David
Welch, Ivo

Publication Date

2018

DOI

10.1057/978-1-349-95189-5_2727

Peer reviewed



Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 168 (2017) 372–399

JOURNAL OF
**Economic
Theory**

www.elsevier.com/locate/jet

Stability with one-sided incomplete information

Sushil Bikhchandani

UCLA Anderson School of Management, Los Angeles, CA 90095-1481, United States

Received 5 October 2015; final version received 6 January 2017; accepted 7 January 2017

Available online 11 January 2017

Abstract

Two notions of stability, ex ante stability and Bayesian stability, are investigated in a matching model with non-transferrable utility, interdependent preferences, and one-sided incomplete information. Ex ante stable matching-outcomes are unblocked for every belief on the blocking partner's type while Bayesian stable matching-outcomes are unblocked with respect to prior beliefs. Ex ante stability is a minimal requirement. Bayesian stability is a more selective desideratum with sound efficiency properties.

© 2017 Elsevier Inc. All rights reserved.

JEL classification: C78; D47; D82

Keywords: Stable matching; Incomplete information; Incomplete-information core

1. Introduction

The concept of stability in two-sided matching markets was introduced in a seminal paper by [Gale and Shapley \(1962\)](#). A matching is stable if there does not exist a blocking pair, i.e., no two agents prefer each other to their respective partners in the matching. In this model, agents do not have any private information that might affect their own preferences or the preferences of other agents. Much of the subsequent literature also assumes complete information – see [Roth and Sotomayor \(1990\)](#).

At least some private information is present in most matching markets. Surgeons at a hospital do not know the operating skill of an applicant to its surgical internship program; this is revealed only after the applicant is hired. A firm has an imprecise estimate about a potential employee's

E-mail address: sbikhcha@anderson.ucla.edu.

productivity, based on his experience and training; the firm learns the worker's productivity only after employment. Similarly, standardized test scores, grades, and reference letters provide only partial information about the academic ability of a college applicant. Thus, it is important to investigate matching markets where these is incomplete information.

In this paper, I explore two notions of stability under one-sided incomplete information in a matching model in which workers are matched with firms.¹ Utility is non-transferrable (NTU) and side payments between workers and firms are not possible. Agents have interdependent preferences in the sense that the utility of an agent depends on the type of the agent it is matched with. Workers have private information about their types whereas firms' types are common knowledge. Once a matching occurs, the type of a worker is revealed to the firm it is matched with. The utility function of each agent increases in his own type and the type of its matched partner.

A *matching outcome* is a matching together with the types of workers.² Because a matching may be stable at one set of worker types but not at another, a matching outcome is the appropriate object of stability. The focus is on restrictions imposed on the set of matching outcomes by two notions of stability under incomplete information: *ex ante stability* and *Bayesian stability*. These are defined by the absence of one of two forms of blocking: *blocks for all admissible worker types* and *Bayesian blocks*, respectively. Stable matching-outcomes, once established, are never blocked by any coalition of workers and firms. How stable matching-outcomes might arise is not addressed in this paper.

Bayesian stability is investigated under the assumption that workers' types are independently and identically distributed. A distributional assumption is not necessary for *ex ante stability*.

Ex ante stability adapts to a NTU model, a definition of stability under one-sided incomplete information introduced by Liu et al. (2014) in a matching model with transferrable utility (TU) and side payments. A firm j participates in a block with a worker i only if j is better off with all *admissible types* of i , i.e., with all types of worker i that are better off in the contemplated block. Under this definition of blocking, it is common knowledge between the pair (i, j) that each is better off in the block. Matching outcomes that are unblocked for all admissible worker-types are *ex ante stable*.

A matching outcome is *Bayesian stable* if it is not *Bayesian blocked*. A firm j participates in a *Bayesian block* with worker i if j 's expected utility in the block is greater than its utility in the current matching. The firm's expectation is taken over all admissible types of its blocking partner (worker i). Thus, *Bayesian blocking* is a weaker requirement than *blocking for all admissible types*. Consequently, *Bayesian stable matching-outcomes* are a subset of *ex ante stable matching-outcomes*.

In each of the two notions of stability, the absence of a block to a matching outcome implies that certain states of nature (i.e., vector of worker types) did not occur. Elimination of these states of nature from consideration opens up the possibility of other potential blocks; if these other possible blocks do not transpire, then the implication is that some other states of nature did not occur. Thus, the continued persistence of a matching outcome leads to a recursive decrease in the set of possible states of nature. If a matching endures, it becomes common knowledge among agents that the types of workers are such that there are no blocking opportunities; such matching outcomes are *stable*. Liu et al. (2014) point out that this is similar in spirit to Holmstrom and Myerson (1983)'s notion of *durable mechanisms*.

¹ While the model is of one-to-one matching, it is easily generalized to a many-to-one matching model under the assumption that each firm has responsive preferences.

² Firms' types are common knowledge and remain fixed throughout the analysis.

In one-to-one matching models with complete information, only two-agent coalitions are essential. With incomplete information, however, incentive-compatible information sharing among members of a larger coalition may create new blocking opportunities. It is established that even with incomplete information the essential coalitions involve two agents.

The set of ex ante stable matching-outcomes, which includes the set of Bayesian stable matching-outcomes and also the set of complete-information stable matching-outcomes, is, in a sense, too large. If all matches are individually rational and preferences are anonymous³ then every maximal matching⁴ in which firms of the highest types are matched is ex ante stable. Observe that maximality and firms with highest types matched are very weak requirements for stability: even before worker types are realized each agent knows that a matching will be blocked for all admissible types if it does not satisfy these two requirements.

Blocking for all admissible types sets a high bar to overturn a matching outcome. As a result, ex ante stability imposes a minimal restriction on matching outcomes. Further, even with a “small” amount of incomplete information, ex ante stable matching-outcomes may not be “close” to complete-information stable matchings. This discontinuity is another drawback of ex ante stability. Thus, ex ante stability is not an appealing solution concept for NTU models with incomplete information.

Bayesian stability is a more selective benchmark. Under anonymous preferences, the set of Bayesian stable matching-outcomes consists of maximal matchings in which firms with higher types are matched to stochastically larger worker types. Consequently, the set of Bayesian stable matching-outcomes is interim efficient. If, in addition to anonymous preferences, agents are ex ante symmetric and match utilities are supermodular, then Bayesian stable matching-outcomes are ex ante efficient. A sufficient condition for interim efficiency of Bayesian stable matching-outcomes is also obtained for the case when preferences are not anonymous. In sum, the set of Bayesian stable matching-outcomes has sound efficiency properties. However, Bayesian stability is not prior-free. Further, while the set of Bayesian stable matching-outcomes is non-empty, there exist worker-type vectors at which no matching is Bayesian stable.

The paper concludes with a centralized mechanism that is ex post incentive compatible under one-sided incomplete information model. In this mechanism, the worker-optimal complete-information stable matching is implemented.

RELATED LITERATURE

There are several papers on matching with incomplete information and interdependent preferences, in addition to Liu et al. (2014). Pomatto (2015) investigates the epistemic foundations of Liu et al.’s concept of incomplete-information stability. In a model with two-sided uncertainty and NTU preferences, Lazarova and Dimitrov (2013) investigate stability under a more permissive notion of blocking than Bayesian blocking: a pair of agents blocks a matching if there is positive probability that each agent will do better. Chakraborty et al. (2010) examine the incomplete-information stability of a mechanism, rather than of a matching, in a college-admissions model. Students’ preferences are known but their quality is unknown; preferences of colleges over students depend on student quality. Chakraborty et al. (2010) show that stable mechanisms do not usually exist. Dizdar and Moldovanu (2013) establish that (under transferrable utility) only fixed-proportion sharing rules are compatible with efficiency.

³ That is, each agent’s utility does *not* depend on the identity of the matched agent.

⁴ In a maximal matching, there may be an unmatched worker or there may be an unmatched firm, but not both.

There is an earlier literature on incomplete-information matching models with privately-known preferences. [Dubins and Freedman \(1981\)](#) and [Roth \(1982\)](#) independently show that in the deferred-acceptance mechanism it is a (weakly) dominant strategy for proposers to truthfully report their preferences. Moreover, every Nash equilibrium of the deferred-acceptance mechanism in which proposers follow their dominant strategy leads to a stable outcome ([Roth, 1984](#)). However, [Gale and Sotomayor \(1985\)](#) show that by misreporting their preferences the non-proposing agents can achieve a stable outcome that is more favorable to them. [Roth \(1982\)](#) establishes that there exists no stable mechanism in which it is a dominant strategy for all agents to truthfully reveal their preferences, while [Roth \(1989\)](#) generalizes this negative result under the weaker incentive constraints of Bayes Nash equilibrium. [Ehlers and Masso \(2007\)](#) show that in a matching model with two-sided incomplete information, an ordinarily Bayesian incentive compatible mechanism exists if and only if there is exactly one stable matching at every state of the world. A related negative result is obtained by [Majumdar \(2003\)](#). [Yenmez \(2013\)](#) investigates the existence of stable, efficient, and budget-balanced mechanisms in a model with transfers.

The literature on the core with incomplete information, surveyed in [Forges et al. \(2002\)](#), is also relevant. [Wilson \(1978\)](#), the first paper in the area, notes that information-sharing assumptions at the interim stage are critical in determining blocks to a potential core allocation; minimal information-sharing within coalitions yields Wilson's concept of the coarse core while maximal information-sharing yields the fine core, which is a subset of the coarse core. [Vohra \(1999\)](#) obtains refinements of Wilson's coarse core and fine core by requiring Bayesian incentive compatibility in information sharing. [Dutta and Vohra \(2005\)](#) propose the credible core, in which members of a blocking coalition draw inferences from the nature of the contemplated objection. While the underlying model in the preceding papers is an exchange economy, [Forges \(2004\)](#) establishes the existence of the coarse core in a matching market.

This paper is organized as follows. The basic model is presented in Section 2. Ex ante stability is investigated in Section 3 and Bayesian stability in Section 4. An incentive compatible mechanism is presented in Section 5. Section 6 concludes. All proofs are in an appendix.

2. The model

There are $i = 1, 2, \dots, n$ workers and $j = 1, 2, \dots, m$ firms. Worker i 's type, w_i , is in the interval $[\underline{w}, \bar{w}]$ while firm j 's type, f_j , is in the interval $[\underline{f}, \bar{f}]$. Let $\mathbf{w} = (w_1, w_2, \dots, w_n)$ and $\mathbf{f} = (f_1, f_2, \dots, f_m)$ be the type vectors of workers and firms, respectively. Each worker's type is his private information, whereas \mathbf{f} is common knowledge among all workers and firms. If firm j and worker i are matched together then their respective utilities are:

$$v_j(w_i, f_j, i) \quad \text{and} \quad u_i(w_i, f_j, j) \quad (1)$$

The utility of firm j , v_j , depends on its own type, f_j , the type of the worker it is matched with, w_i , and the worker's identity, i . The type of the worker may represent the productivity of the worker, which is unknown to the firm at the time the worker is hired. There may be other characteristics of the worker, such as education and experience, that are observable to the firm before the worker is hired and these are captured by the dependence of v_j on i . Similarly, the utility of worker i , u_i , depends on the worker's own type and the type and the identity of the firm it is matched with. All the characteristics of a firm are common knowledge.

If worker i and firm j are matched, then firm j learns worker i 's type, w_i . For example, if w_i is worker i 's productivity then the firm learns w_i after the worker is hired. This assumption plays an important role in the analysis.

A dummy worker is indexed $i = 0$ and a dummy firm is indexed $j = 0$, each with type \emptyset . An unmatched worker (firm) is matched to the dummy firm (worker). The utility of an unmatched firm or worker is normalized to zero:

$$v_j(\emptyset, f_j, 0) = 0 \quad \text{and} \quad u_i(w_i, \emptyset, 0) = 0.$$

There are no side payments between matched workers and firms in this model. This may appear counter-factual as firms pay wages to workers. The important assumption is that if side payments are present then there is a standard payment (wage) over which there is little or no bargaining. This is the case with medical residents, law clerks, or college interns who view the job as building their human capital (see Roth and Sotomayor, 1990, p. 125). If a firm makes the same payment to any worker it might hire, the side payment need not be explicitly modeled and is reflected in the utilities of the matched firm and worker. The model here may also be appropriate for matching students to schools.

The focus of this paper is a matching model with one-sided incomplete information. Typically, firms and colleges have a longer history than workers and students. Consequently, there is a great deal of publicly available information about firms and colleges. Therefore, one-sided incomplete information captures an essential element of these environments.

A matching is a function $\mu : \{1, 2, \dots, n\} \rightarrow \{0, 1, 2, \dots, m\}$, where $\mu(i)$ is the firm that worker i is matched with. If $\mu(i) = 0$ then worker i is unmatched and if $\mu(i) = \mu(i')$, $i \neq i'$ then $\mu(i) = \mu(i') = 0$. It is notationally convenient to define the inverse matching function ν of μ , where for any $j \in \{1, 2, \dots, m\}$,⁵

$$\nu(j) \equiv \mu^{-1}(j) = \begin{cases} i, & \text{if there is an } i \in \{1, 2, \dots, n\} \text{ s.t. } \mu(i) = j, \\ 0, & \text{otherwise.} \end{cases}$$

A matching may be stable for some agent types but not for other. Thus, the appropriate object for defining stability is a matching together with the types of workers and firms. Recall that $\mathbf{w} = (w_1, w_2, \dots, w_n)$ is the type vector for workers. A *matching outcome*, (μ, \mathbf{w}) , is a matching function together with a worker type vector.⁶ A matching outcome (μ, \mathbf{w}) is *individually rational* if

$$\begin{aligned} u_i(w_i, f_{\mu(i)}, \mu(i)) &\geq 0, & \forall i \\ v_j(w_{\nu(j)}, f_j, \nu(j)) &\geq 0, & \forall j \end{aligned}$$

Let Σ^0 be the set of individually-rational matching-outcomes. If utility is increasing in worker type, then $(\mu, \mathbf{w}) \in \Sigma^0$ implies $(\mu, \hat{\mathbf{w}}) \in \Sigma^0$ for all $\hat{\mathbf{w}} \geq \mathbf{w}$.

The following assumptions on agents' utility functions are invoked in some of the propositions.

INCREASING UTILITY: *The utility functions $u_i(w_i, f_j, j)$ and $v_j(w_i, f_j, i)$ are strictly increasing in w_i and f_j , for all i and j .*

The assumption of increasing utility functions is important for most of the results.⁷

⁵ Throughout the paper, ν represents the inverse matching of μ , $\hat{\nu}$ represents the inverse matching of $\hat{\mu}$, etc.

⁶ Because the vector of firm types, \mathbf{f} , is common knowledge among all agents and fixed throughout, it is dropped from the notation of a matching outcome for brevity.

⁷ Lemmas 1 and 2 do not assume increasing utility. Propositions 2, 3, 4, 5 and 9 assume that only firms' utility functions are increasing.

ANONYMOUS PREFERENCES: Preferences are anonymous if utilities depend only on the type and not on the identity of the matched agent. That is, worker-firm pair (i, j) is matched to each other then their utilities are:

$$v_j(w_i, f_j) \text{ and } u_i(w_i, f_j)$$

If, for instance, two workers i and i' have the same type, $w_i = w_{i'}$, then firm j derives the same utility from matching with i or i' . Anonymous preferences is a plausible assumption when the type of an agent is a primary determinant of the utility of a match.

ALL MATCHES ARE INDIVIDUALLY RATIONAL:

$$u_i(\underline{w}, f_j, j) \geq 0, v_j(\underline{w}, f_j, i) \geq 0, \text{ for all } i, j.$$

Before investigating stability under incomplete information, it is useful to recall the definition of complete-information stability. An individually-rational matching-outcome $(\mu, \mathbf{w}) \in \Sigma^0$ is complete-information blocked if there is a worker-firm pair (i, j) such that

$$u_i(w_i, f_j, j) > u_i(w_i, f_{\mu(i)}, \mu(i)) \text{ and } v_j(w_i, f_j, i) > v_j(w_{v(j)}, f_j, v(j)) \tag{2}$$

If there does not exist an (i, j) satisfying (2), then (μ, \mathbf{w}) is complete-information stable. Gale and Shapley (1962) provide a constructive proof of existence of a complete-information stable matching; their deferred-acceptance algorithm stops at a stable matching.

Next, two notions of stability under incomplete information are developed, the first of which is less restrictive than the second.

3. Ex ante stability

In this section, I present a definition of blocking and stability that adapts, to a model without side payments, the definition of Liu et al. (2014).

Under complete information, a matching is blocked if there is a worker and a firm that can each improve its utility by matching with each other rather than with their respective partners in the matching. Moreover, it is common knowledge that each is better off in the block. The definition of blocking under incomplete information given below satisfies a similar requirement.

Let A be a set of matching outcomes with ‘ μ in A ,’ i.e., there exists $\hat{\mathbf{w}}$ such that $(\mu, \hat{\mathbf{w}}) \in A$. Suppose that worker i and firm j are considering a block to μ . The set of admissible worker-types for this block are

$$A^{ij}(\mu) \equiv \{ \mathbf{w}' \mid (\mu, \mathbf{w}') \in A, u_i(w'_i, f_j, j) > u_i(w'_i, f_{\mu(i)}, \mu(i)) \} \tag{3}$$

These are the set of worker type vectors at which worker i prefers firm j to its current match $\mu(i)$.

A matching-outcome $(\mu, \mathbf{w}) \in A \subseteq \Sigma^0$, where $\mathbf{w} = (w_1, w_2, \dots, w_n)$, is A -blocked for all admissible worker-types if there is a worker-firm pair (i, j) satisfying

$$\mathbf{w} \in A^{ij}(\mu) \tag{4}$$

$$\text{and for all } \mathbf{w}' \in A^{ij}(\mu) \text{ s.t. } w'_{v(j)} = w_{v(j)},$$

$$v_j(w'_i, f_j, i) > v_j(w_{v(j)}, f_j, v(j)) \tag{5}$$

Condition (4) requires that worker i is better off in the potential block. The corresponding condition for firm j is as follows. Firm j does not know worker i 's type. Therefore, in order to

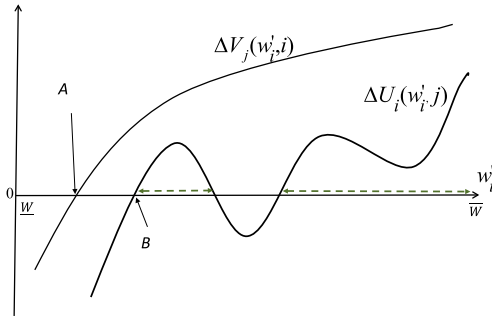


Fig. 1. Blocking by (i, j) is possible.

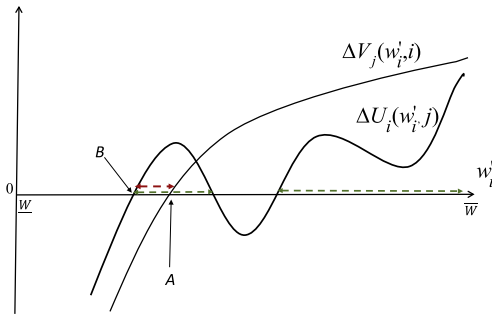


Fig. 2. Blocking by (i, j) is impossible.

participate in the block, firm j should be better off with any worker i type w'_i that would benefit from the block; that is, for any $w' \in A^{ij}(\mu)$ that is consistent with firm j 's knowledge about worker $v(j)$.⁸ If (4) and (5) are satisfied then the pair (i, j) blocks (μ, w) in A for all admissible worker types. Under this definition, the fact that worker i and firm j are willing to participate in a block makes it common knowledge that each is better off in the block.

Figs. 1 and 2 illustrate this definition of blocking. For simplicity, assume that $(\mu, w) \in A$ for any $w \in [\underline{w}, \bar{w}]^n$. In both figures, the gain to worker i in switching from firm $\mu(i)$ to firm j as a function of w'_i is⁹

$$\Delta U_i(w'_i, j) = u_i(w'_i, f_j, j) - u_i(w'_i, f_{\mu(i)}, \mu(i))$$

The two intervals indicated by the (green) broken line-segments is the set of w'_i that satisfy $u_i(w'_i, f_j, j) > u_i(w'_i, f_{\mu(i)}, \mu(i))$, i.e., the set of admissible worker i types. The smallest intersection of $\Delta U_i(w'_i, j)$ with the horizontal axis is from below; this point is labeled B . The gain to firm j in switching from worker $v(j)$ to worker i is

$$\Delta V_j(w'_i, i) = v_j(w'_i, f_j, i) - v_j(w_{v(j)}, f_j, v(j)), \tag{6}$$

⁸ This is the requirement that $w'_{v(j)} = w_{v(j)}$. If $v(j) = 0$, then this condition has no effect.

⁹ The dependence of ΔU_i on f_j and $\mu(i)$ is suppressed in the notation. A similar comment applies to ΔV_j defined below.

which is an increasing function of w'_i if firm j 's utility is increasing in the type of the matched worker. It intersects the horizontal axis at A . The difference in the two figures is in the relative location of the points A and B .

In Fig. 1, B lies to the right of A . Therefore, each admissible type of worker i is preferred by firm j to its current match $v(j)$. Hence, matching μ is A -blocked for all admissible types¹⁰ by worker i and firm j whenever w_i is in one of the two intervals indicated by (green) broken line-segments.

In Fig. 2, on the other hand, B is to the left of A . There are admissible worker i types (those in the interval indicated by a (red) broken thick line-segment between B and A) that are not preferred by firm j to its current match $v(j)$. Thus, the matching μ is never A -blocked for all admissible types by worker i and firm j .

If a matching outcome is A -blocked for all admissible types, then it is blocked for *any* “reasonable belief” over worker types. Here, reasonable belief means that firms do not entertain the possibility that a rational worker would participate in a block that makes the worker worse off. That is, only $w' \in A^{ij}(\mu)$ are considered.

Conversely, if a matching μ is not blocked then agents should infer that worker types w at which (μ, w) is A -blocked for all admissible types did not occur; for each worker-firm pair (i, j) that are not matched to each other at μ , these are values of worker $v(j)$ types that yield Fig. 1 together with worker i types that are in the two broken line-segments in Fig. 1. These “never stable” matching-outcomes should be eliminated from consideration.¹¹ The consequent reduction in the set of matching outcomes associated with μ opens up additional possibilities of blocking. Successive elimination of never-stable matching-outcomes leads to a stable set of matching outcomes. This process is reminiscent of the notion of rationalizability in game theory due to [Bernheim \(1984\)](#) and [Pearce \(1984\)](#), where “never-best response” strategies are successively eliminated.

A matching outcome $(\mu, w) \in A$ is *ex ante A-stable* if it is not A -blocked for all admissible types. The term *ex ante* is used because the stability of an unblocked matching outcome is ascertained before (or without) assessing a probability distribution over worker types. The set A is *ex ante self-stabilizing* if every $(\mu, w) \in A$ is *ex ante A-stable*.

If the set $\{(\mu, w)\}$ is *ex ante self-stabilizing* then, by definition, μ is a complete-information stable matching at w . Additional results are gathered in [Lemmas 1 and 2](#) below; similar results are obtained in [Liu et al. \(2014\)](#) for a model with side payments.

Lemma 1.

- (i) Suppose that $B \subset A$, where A and B are sets of matching outcomes. If $(\mu, w) \in B$ is *ex ante B-stable* then it is *ex ante A-stable*.
- (ii) Let μ_w be a complete-information stable matching at w . Then, (μ_w, w) is *ex ante A-stable* for any A such that $(\mu_w, w) \in A$.

Next, following [Liu et al. \(2014\)](#), an *ex ante stable set* is defined by iterative elimination of blocked matching-outcomes. Successive elimination of matching outcomes that are blocked for all admissible types leads to an *ex ante stable set*.

¹⁰ Hereafter, I use this shorter form for brevity.

¹¹ I am grateful to an anonymous referee for this interpretation of blocked matching outcomes.

Recall that Σ^0 is the set of individually-rational matching-outcomes. Define

$$\Sigma^k = \{(\mu, \mathbf{w}) \in \Sigma^{k-1} \mid (\mu, \mathbf{w}) \text{ is not } \Sigma^{k-1}\text{-blocked for all admissible types}\} \tag{7}$$

Then

$$\Sigma^* \equiv \bigcap_{k=0}^{\infty} \Sigma^k$$

is the set of *ex ante stable matching-outcomes*.

For any \mathbf{w} , let $\mu_{\mathbf{w}}$ be a matching that is complete-information stable matching at \mathbf{w} ; we know from Gale and Shapley (1962) that for each \mathbf{w} there exists such a $\mu_{\mathbf{w}}$. As a complete-information stable is individually rational, we have $\{\mu_{\mathbf{w}}, \mathbf{w}\} \in \Sigma^0$. By repeated application of Lemma 1(ii) it follows that $\{\mu_{\mathbf{w}}, \mathbf{w}\} \in \Sigma^k$, for all k and hence Σ^* is non-empty.

It is clear from (7) that $\Sigma^k \subseteq \Sigma^{k-1}$ and that Σ^* is self-stabilizing. The next result shows that Σ^* is the largest self-stabilizing set.

Lemma 2. *If A is an ex ante self-stabilizing set then $A \subseteq \Sigma^*$.*

Blocking for all admissible types may appear to be too stringent a requirement. Therefore, the associated set of stable matching outcomes, Σ^* , may be rather large. This is confirmed in the next section.

3.1. Anonymous preferences

Under anonymous preferences, the set of ex ante stable matching-outcomes is described by two characteristics, defined below.

A matching μ is a *maximal matching* if it has the largest possible number of matched worker-firm pairs. Thus, all agents on the shorter side of the market are matched at a maximal matching. Maximal matchings need not be individually rational, unless all matches are individually rational.

Firms of highest types are matched at μ if $v(j) \neq 0$ implies $v(\hat{j}) \neq 0$ for any \hat{j} such that $f_{\hat{j}} > f_j$.

Proposition 1 shows that any maximal matching in which firms of the highest types are matched is in the set Σ^* for almost all \mathbf{w} . Maximality and highest types of firms matched is a minimal requirement for stability; under the assumptions of the proposition, any matching that does not satisfy either of these two requirements is Σ^0 -blocked.

Proposition 1. *Assume that utility is increasing, agents have anonymous preferences, and all matchings are individually rational. Then*

- (i) *If μ is a maximal matching at which firms of the highest types are matched then for any $\mathbf{w} > \underline{\mathbf{w}}$, $(\mu, \mathbf{w}) \in \Sigma^*$.*¹²
- (ii) *If matching μ is either not maximal or firms of the highest types are not matched then for any $\mathbf{w} > \underline{\mathbf{w}}$, (μ, \mathbf{w}) is Σ^0 -blocked for all admissible types.*

¹² Note that $\mathbf{w} > \underline{\mathbf{w}}$ [$\equiv (\underline{w}_1, \underline{w}_2, \dots, \underline{w}_n)$] is the statement that $w_i > \underline{w}_i$ for each worker i .

Thus, under the assumptions of [Proposition 1](#), ex ante stability provides little restriction on matching outcomes. This is illustrated starkly in the following example.

Example 1 (Assortative Matching). Let the number of workers equal the number of firms, $n = m$. The firm types are $f_j = j$. The utility functions of worker i and firm j , when matched to each other, are

$$u_i(w_i, j) = v_j(w_i, j) = jw_i$$

Then, with $0 < w_{i_1} < w_{i_2} < \dots < w_{i_n}$, the unique complete-information stable-matching pairs worker i_j with firm j , the positive assortative matching. But [Proposition 1](#) implies that all maximal matchings are ex ante stable for almost all \mathbf{w} .

To see this directly, consider the negative assortative matching where worker i_n is matched to firm 1, worker i_{n-1} is matched to firm 2, etc. Firm n is matched to the lowest type worker i_1 but it does not know that i_1 has the lowest type; firm n will reject a blocking proposal from all other workers, including worker i_n . This is because worker i_n types $w'_{i_n} < w_{i_1}$ would also do better by matching with firm n than with their current match firm 1.

Similarly, every firm $j \geq 2$ will reject a blocking proposal from any other worker. No worker will propose a block with firm 1. \square

Remark 1(i). In this example, the positive assortative match is the only stable matching under complete information. However, if there is a little one-sided incomplete information, i.e. $\bar{w} = \underline{w} + \epsilon$ for any $\epsilon > 0$, then the set of ex ante stable matching-outcomes consists of all maximal matching-outcomes including the negative assortative match. Because a small amount of incomplete information is present in most matching markets, this discontinuity in the set of stable matching-outcomes is a drawback of the concept of ex ante stability.

Remark 1(ii). In a TU model with side payments, Proposition 3 of [Liu et al. \(2014\)](#) implies that in any positive-assortative matching model, only the positive-assortative matching-outcome is ex ante stable. In an NTU model, however, ex ante stability has little predictive power, as the above example and [Proposition 1](#) demonstrate. This further highlights the important role in information revelation played by the availability of side payments in [Liu et al. \(2014\)](#).

Ex ante stability provides some restriction when one relaxes the assumption of anonymous preferences for all agents. However, the set of stable matching-outcomes remains large. Consider, for example, the case when workers have anonymous preferences and firms have any non-anonymous preferences with utility increasing in worker type. Let $\underline{\mu}$ be a complete-information stable matching at the lowest possible worker-type vector $\underline{\mathbf{w}}$. Then $(\underline{\mu}, \mathbf{w})$ is an ex ante stable matching-outcome at any $\mathbf{w} \geq \underline{\mathbf{w}}$.

Next, it is shown that blocking coalitions of size greater than two provide no additional restriction on the set of ex ante stable matching-outcomes. This is true whether or not preferences are anonymous and whether or not all matchings are individually rational.

3.2. Coalitions with multiple agents

In a complete-information model, if a matching is not blocked by a worker-firm pair then it is not blocked by a larger coalition. In an incomplete-information setting, however, it is conceivable

that a larger coalition might block a matching that is unblocked by any two-agent coalition. It is shown that this is not the case.

As in two-person blocking, assume that firms draw inferences about worker types only from the membership of a (larger) blocking coalition. The pairing of worker i and firm j within a larger coalition to block matching μ implies that $\mathbf{w} \in A^{ij}(\mu)$. This has two implications. First, as in a two-person blocking coalition, firm j knows that w_i is such that worker i prefers firm j to firm $\mu(i)$. Second, if $v(j)$, the worker matched with firm j under μ , is also a member of the larger coalition, then the inference about $w_{v(j)}$ is that its value could not be too high, else firm j would not have joined the coalition; this is useful to the firm paired with worker $v(j)$ in the blocking coalition. The argument is sketched out informally below.

Consider a matching μ and suppose that (i, j) are willing to block μ as part of a coalition S which has more than two members. Suppose also that worker $v(j)$, $v(j) \neq i$, is a member of S . Let $\hat{j} \in S$, $\hat{j} \neq j$, be the firm that is to be matched with $v(j)$ in S . What can firm \hat{j} conclude about $w_{v(j)}$? If firm j 's utility function is increasing (in the type of the worker it is matched to) then from (6) it is clear that ΔV_j is decreasing in $w_{v(j)}$. The only difference between Figs. 1 and 2 is that worker $v(j)$'s type is greater in Fig. 2 and hence firm j will never block with worker i . Let $w_{v(j)}^*$ be the value of worker $v(j)$'s type that represents the transition point between Figs. 1 and 2. That is, any $w_{v(j)} < w_{v(j)}^*$ would yield a picture like Fig. 1 and any $w_{v(j)} > w_{v(j)}^*$ would give us Fig. 2. Thus, because (i, j) are willing to block μ , firm \hat{j} can conclude that $w_{v(j)} \leq w_{v(j)}^*$.

Any blocking coalition in which each member is strictly better off consists of worker-firm pairs; an unmatched worker (or firm) in the coalition cannot be strictly better off in the block than in the individually-rational matching-outcome being blocked. Thus, let $(i_1, j_1), \dots, (i_k, j_k)$ be a blocking coalition, where each pair (i_ℓ, j_ℓ) , $\ell = 1, 2, \dots, k$ is matched together in the proposed blocking coalition.

Suppose that firms' utility functions are increasing. Let A be a set of individually-rational matching-outcomes. Then $(\mu, \mathbf{w}) \in A \subseteq \Sigma^0$ is A -blocked for all admissible worker types by a coalition $(i_1, j_1), \dots, (i_k, j_k)$ if for all $\ell = 1, \dots, k$ ¹³

$$\mathbf{w} \in A^{i_\ell j_\ell}(\mu) \tag{8}$$

and for all $\mathbf{w}' \in A^{i_\ell j_\ell}(\mu)$ s.t. $w'_{v(j_\ell)} = w_{v(j_\ell)}$, $w'_{i_\ell} \leq w_{i_\ell}^*$,

$$v_{j_\ell}(w'_{i_\ell}, f_{j_\ell}, i_\ell) > v_{j_\ell}(w_{v(j_\ell)}, f_{j_\ell}, v(j_\ell)) \tag{9}$$

This definition is similar to the earlier definition of blocking with a two-member coalition except for the constraints $w'_{i_\ell} \leq w_{i_\ell}^*$. If, say, worker $i_1 \neq v(j_\ell)$ for any $\ell \geq 2$ then $w_{i_1}^* = \bar{w}$. Thus, (8) and (9) are identical to (4) and (5) for the two-person coalition (i_1, j_1) . If, instead, $i_1 = v(j_\ell)$ for some $\ell \geq 2$ then, as argued above, an upper bound of $w_{i_1}^*$ on worker i_1 's type is implied. However, note that if firms' utility functions are increasing, then $w'_{i_1} \leq w_{i_1}^*$ is redundant: if (9) is satisfied for w'_{i_1} then it is satisfied for all $w''_{i_1} > w'_{i_1}$. Hence, the following result is immediate:

Proposition 2. *Suppose that firms' utility functions are increasing. If a matching outcome is A-blocked for all admissible types by a coalition with four or more members, then each matched worker-firm pair within the coalition A-blocks for all admissible types as a two-member coalition.*

¹³ See (3) for a definition of $A^{i_\ell j_\ell}(\mu)$.

Thus, allowing for blocks by larger coalitions does not reduce the size of Σ^* , the largest self-stabilizing set. The conclusion of Section 3.1, that ex ante stability has little predictive power, especially when workers have anonymous preferences, is not altered when larger coalitions are considered. Bayesian stability, a more discerning notion, is considered next.

4. Bayesian stability

In markets which meet repeatedly, such as colleges-students or hospitals-interns, history facilitates the formation of a probability distribution over the types of the next cohort of “workers.” Therefore, Bayesian stability is a natural notion for such environments. Assume that it is common knowledge that each worker’s type is independently and identically distributed with cumulative probability distribution function F . The rest of the model is as described in Section 2.

A matching μ is Bayesian blocked by a worker-firm pair, (i, j) , if worker i ’s utility increases and firm j ’s *expected* utility increases in the block. After (i, j) block the matching μ , firm j will learn w_i . It may turn out that w_i is small enough that firm j is worse off in the block, i.e., $v_j(w_i, f_j, i) < v_j(w_{v(j)}, f_j, v(j))$. A definition of Bayesian blocking depends on whether firm j may return to worker $v(j)$, perhaps at some cost.¹⁴ I assume that firm j cannot return to the worker to which it was previously matched. This is an accurate description of some settings. For instance, a student who is asked to leave a Ph.D. program, or departs to another graduate school, seldom returns.

In other environments, a previously-matched pair may be matched again in the future, at some cost. A definition of Bayesian blocking with costly return to a status quo matching is provided in Section 4.3. The focus in this paper is on the case where return to a previous match is not allowed or is prohibitively costly. The efficiency results below provide a benchmark for settings where return at some cost is possible.

Let A be a set of individually-rational matching-outcomes, where individual rationality is as defined in Section 2. Then $(\mu, \mathbf{w}) \in A$ is *Bayesian A-blocked* if there is a worker-firm pair (i, j) satisfying

$$\mathbf{w} \in A^{ij}(\mu) \tag{10}$$

$$E[v_j(w'_i, f_j, i) | \mathbf{w}' \in A^{ij}(\mu), w'_{v(j)} = w_{v(j)}] > v_j(w_{v(j)}, f_j, v(j)) \tag{11}$$

The set of admissible worker types, $A^{ij}(\mu)$, is defined in (3). Condition (10) requires that worker i is better off in the potential block. Inequality (11) requires that firm j ’s expected utility in the block is greater than its current utility. The conditional expectation in (11) is over \mathbf{w}' in the set $A^{ij}(\mu)$ that are consistent with j ’s knowledge of worker $v(j)$ ’s type. The expectation is taken with the i.i.d. probability distributions F over worker types.

In Figs. 1 and 2, the projection of the set $A^{ij}(\mu)$ on to worker i ’s type-space consists of the two intervals indicated by (green) broken line-segments. Bayesian blocking is possible even in Fig. 2 if the probability of worker i ’s type being in the (red) thick line-segment, where $v_j(w'_i, f_j, i) < v_j(w_{v(j)}, f_j, v(j))$, is relatively small. As noted above, if worker i ’s type is in this region where firm j is worse off, then firm j would like to return but cannot. The expected utility on the left-hand side of (11) is based on the assumption that firm j cannot go back to worker $v(j)$.

Recall that a matching outcome is ex ante stable if it is unblocked for every belief on the set of admissible types. The requirement for Bayesian stability is weaker: that a matching outcome be

¹⁴ Worker i will not regret the decision to block as the type of firm j is common knowledge.

unblocked with respect to the prior belief on the set of admissible types. An individually-rational matching-outcome $(\mu, \mathbf{w}) \in A$ is *Bayesian A-stable* if it is not Bayesian A-blocked. A set A is *Bayesian self-stabilizing* if every $(\mu, \mathbf{w}) \in A$ is Bayesian A-stable.

Let $\Gamma^0 (= \Sigma^0)$ be the set of individually-rational matching-outcomes. For $k \geq 1$ define

$$\Gamma^k = \{(\mu, \mathbf{w}) \in \Gamma^{k-1} \mid (\mu, \mathbf{w}) \text{ is not Bayesian } \Gamma^{k-1}\text{-blocked}\}$$

For each matching μ , the probability distribution over a worker’s type is updated at each stage k by eliminating types that, together with μ , would be Bayesian Γ^{k-1} -blocked. The set of *Bayesian stable matching-outcomes* is

$$\Gamma^* = \bigcap_{k=0}^{\infty} \Gamma^k$$

Observe that $\Gamma^k \subseteq \Gamma^{k-1}$, by definition. Further, Γ^* is Bayesian self-stabilizing and, as shown next, non-empty.

Proposition 3. *If firms have increasing utility then $\Gamma^* \neq \emptyset$.*

However, as the following example shows, there may exist \mathbf{w} such that $(\mu, \mathbf{w}) \notin \Gamma^*$ for any μ .

Example 2 *(There exists \mathbf{w} such that $(\mu, \mathbf{w}) \notin \Gamma^*$ for any μ).*

There are two workers and two firms. The utility functions of workers and firms are $u_i(w_i, f_j, j) = v_j(w_i, f_j, i) = jw_i$, $i = 1, 2$, $j = 1, 2$. The workers types are i.i.d. uniform on $[0, 1]$.

Take any $\mathbf{w} = (w_1, w_2) \in (0, 0.5)^2$. It is shown that there does not exist μ such that $(\mu, \mathbf{w}) \in \Gamma^1$. Note that only μ that are maximal need be considered, as any non-maximal matching is Bayesian Γ^0 -blocked. As $\mathbf{w} \in (0, 0.5)^2$, at any μ firm 2 is matched with a worker with type less than 0.5. Therefore, firm 2 and the worker who is matched with firm 1 form a Bayesian Γ^0 -blocking pair as firm 2’s expected value in the block is greater than 0.5. Thus, $(\mu, \mathbf{w}) \notin \Gamma^1$ and there is no μ such that $(\mu, \mathbf{w}) \in \Gamma^*$. \square

For any \mathbf{w} , there exists a complete-information stable matching, $\mu_{\mathbf{w}}$. Thus, the set $\{(\mu_{\mathbf{w}}, \mathbf{w})\}$ is Bayesian self-stabilizing and, as $(\mu_{\mathbf{w}}, \mathbf{w})$ is individually rational, $(\mu_{\mathbf{w}}, \mathbf{w}) \in \Gamma^0$. In **Example 2**, if $\mathbf{w} \in (0, 0.5)^2$ then $(\mu_{\mathbf{w}}, \mathbf{w}) \in \Gamma^0$ but $(\mu_{\mathbf{w}}, \mathbf{w}) \notin \Gamma^1$. Hence, **Lemma 1** does not hold for Bayesian stability. This also shows that Γ^* , which is a Bayesian self-stabilizing set, need not include every other Bayesian self-stabilizing set. Thus, **Lemma 2** also does not extend to Bayesian stability.

Next, it is shown that **Proposition 2** extends to Bayesian blocking; thus, only two-person blocking coalitions need to be considered.

An individually-rational matching-outcome $(\mu, \mathbf{w}) \in A$ is *Bayesian A-blocked by a coalition* $S = \{(i_1, j_1), \dots, (i_k, j_k)\}$ if for all $\ell = 1, \dots, k$,

$$\begin{aligned} & \mathbf{w} \in A^{i_\ell j_\ell}(\mu) \\ & \mathbb{E} \left[v_{j_\ell}(w'_{i_\ell}, f_{j_\ell}, i_\ell) \mid \mathbf{w}' \in A^{i_\ell j_\ell}(\mu), w'_{v(j_\ell)} = w_{v(j_\ell)}, w'_{i_\ell} \leq w^\circ_{i_\ell} \right] \\ & > v_{j_\ell}(w_{v(j_\ell)}, f_{j_\ell}, v(j_\ell)) \end{aligned}$$

The rationale for the restriction $w_{i_\ell} \leq w_{i_\ell}^\circ$ is similar to that in the definition of multiple-agent coalition blocking for all admissible types in Section 3.2. If no worker and firm in S are matched to each other under μ , then $w_{i_\ell}^\circ = \bar{w}$ and this restriction has no effect. If, instead, $i_\ell = v(j_{\hat{\ell}})$ for some $\ell, \hat{\ell} \in \{1, \dots, k\}, \ell \neq \hat{\ell}$, then $w_{i_\ell}^\circ \leq \bar{w}$.¹⁵ In either case, we have

$$\begin{aligned} & \mathbb{E}[v_{j_\ell}(w'_{i_\ell}, f_{j_\ell}, i_\ell) \mid \mathbf{w}' \in A^{i_\ell j_\ell}(\mu), w'_{v(j_\ell)} = w_{v(j_\ell)}] \\ & \geq \mathbb{E}\left[v_{j_\ell}(w'_{i_\ell}, f_{j_\ell}, i_\ell) \mid \mathbf{w}' \in A^{i_\ell j_\ell}(\mu), w'_{v(j_\ell)} = w_{v(j_\ell)}, w'_{i_\ell} \leq w_{i_\ell}^\circ\right] \end{aligned}$$

where the inequality follows from the monotonicity of v_{j_ℓ} in w'_{i_ℓ} . Hence, the following proposition is immediate:

Proposition 4. *Suppose that firms’ utility functions are increasing. If a matching outcome is Bayesian A-blocked by a coalition with four or more members, then each matched worker-firm pair within the coalition constitutes a Bayesian A-blocking pair.*

Thus, as in the case of blocking for all admissible types, it is enough to consider two-person Bayesian blocking coalitions. This facilitates a comparison of blocking and stability under the two notions of blocking presented in the next proposition.

Proposition 5. *Suppose that firms’ utility functions are increasing. Let A, B be two sets of matching outcomes with $B \subseteq A$. If $(\mu, \mathbf{w}) \in B$ is not Bayesian blocked in B , then it is not blocked for all admissible types in A . Consequently, $\Gamma^* \subseteq \Sigma^*$.*

Hence, the set of Bayesian stable matching-outcomes is no larger than the set of ex ante stable matching-outcomes. This raises the question whether Bayesian self-stabilizing sets are, in some sense, efficient. To address this, two notions of efficiency are presented below.

EFFICIENCY

The following notions of efficiency are due to [Holmstrom and Myerson \(1983\)](#). In the definitions below, neither μ nor $\hat{\mu}$ is assumed to be stable.

Consider a stage at which it is common knowledge that worker types are in a set $B \subseteq [\underline{w}, \bar{w}]^n$. A matching $\hat{\mu}$ ex ante dominates in B another matching μ if

$$\begin{aligned} & \mathbb{E}\left[\sum_{i=1}^n u_i(w_i, f_{\hat{\mu}(i)}, \hat{\mu}(i)) + \sum_{j=1}^m v_j(w_{\hat{v}(j)}, f_j, v(j)) \mid \mathbf{w} \in B\right] \\ & > \mathbb{E}\left[\sum_{i=1}^n u_i(w_i, f_{\mu(i)}, \mu(i)) + \sum_{j=1}^m v_j(w_{v(j)}, f_j, v(j)) \mid \mathbf{w} \in B\right] \end{aligned}$$

Matching μ is ex ante efficient in B if it is not ex ante dominated in B . Thus, an ex ante efficient matching maximizes the sum of expected utilities of agents. It implies [Holmstrom and Myerson \(1983\)](#)’s definition of ex ante efficiency, which maximizes a weighted sum of expected utilities.

¹⁵ If $w_{i_\ell}^\circ < \bar{w}$ and worker i_ℓ ’s type is $w_{i_\ell}^\circ$ then firm $j_{\hat{\ell}}$ is indifferent between participating in the blocking coalition or not participating.

A weaker notion is *interim efficiency*. Consider an interim stage at which each worker knows his own type and each matched firm knows the type of the worker to whom it is matched. In addition, it is common knowledge that worker types are in $B \subseteq [\underline{w}, \bar{w}]^n$. Then a matching $\hat{\mu}$ *interim dominates in B* another matching μ if

$$u_i(w_i, f_{\hat{\mu}(i)}, \hat{\mu}(i)) \geq u_i(w_i, f_{\mu(i)}, \mu(i)), \quad \forall \mathbf{w} \in B, \forall i,$$

and $E \left[v_j(w_{\hat{v}(j)}, f_j, \hat{v}(j)) \mid \mathbf{w} \in B \right] \geq v_j(w_{v(j)}, f_j, v(j)), \quad \forall \mathbf{w} \in B, \forall j$

and the inequality is strict for at least one agent at a $\mathbf{w} \in B$. Matching μ is *interim efficient in B* if it is not interim dominated in B .

As [Holmstrom and Myerson \(1983\)](#) note, if a matching is ex ante efficient then it is also interim efficient.

4.1. Anonymous preferences

The set Γ^* has strong efficiency properties under the assumption of anonymous preferences. In order to establish this, the set

$$\Gamma^*(\mu) \equiv \{\mathbf{w} \mid (\mu, \mathbf{w}) \in \Gamma^*\}$$

is characterized in [Proposition 6](#). To simplify the exposition, the assumption that all matchings are individually rational is invoked in this section.

Proposition 6. *Assume that utility functions are increasing, agents have anonymous preferences, and that all matchings are individually rational. Then*

- (i) *A matching outcome $(\mu, \mathbf{w}) \in \Gamma^*$ for some \mathbf{w} if and only if μ is a maximal matching in which firms of the highest types are matched.*
- (ii) *Without loss of generality, assume that $f_1 \geq f_2 \geq \dots \geq f_m$. For the matched firms, $j = 1, 2, \dots, \min[m, n]$, there exist unique w_j^* , with $w_{j-1}^* \geq w_j^*$, such that*

$$\Gamma^*(\mu) = \{\mathbf{w} \mid w_{v(j)} \geq w_j^*, \forall j \leq \min[m, n], \text{ and } w_i \geq \underline{w}, \forall i \text{ s.t. } \mu(i) = 0\}.$$

Further, $w_{j-1}^ > w_j^*$ if and only if $f_{j-1} > f_j$.*

The characterization of μ in Γ^* in [Proposition 6\(i\)](#) is similar to that of μ in Σ^* in [Proposition 1\(i\)](#).¹⁶ The distinguishing feature of Bayesian blocking is [Proposition 6\(ii\)](#), which states that firms with higher types are matched to stochastically larger worker types in $\Gamma^*(\mu)$. As shown below, the efficiency properties of Γ^* flow from this attribute.

Before establishing conditions under which Γ^* is ex ante efficient, the following definitions are needed.

Agents are *ex ante symmetric* if all workers have the same utility function and all firms have the same utility function. That is, there exists a worker utility function $u(w, f, j)$ and a firm utility function $v(w, f, j)$ such that¹⁷

¹⁶ Recall that ‘ μ in A ,’ where A is a set of matching outcomes, is the statement that there exists a \mathbf{w} such that $(\mu, \mathbf{w}) \in A$.

¹⁷ If, in addition, preferences are anonymous, then $u(w, f, j) \equiv u(w, f)$ and $v(w, f, i) \equiv v(w, f)$.

$$u_i(w, f, j) = u(w, f, j), \text{ and } v_j(w, f, i) = v(w, f, i), \quad \forall w, f, i, j$$

The match utility of a worker-firm pair (i, j) is

$$M_{ij}(w, f) = u_i(w, f) + v_j(w, f)$$

The match utility of (i, j) is *supermodular* if $M_{ij}(w, f)$ is (weakly) supermodular in w and f .

Proposition 7. *Assume that utility functions are increasing, agents are ex ante symmetric and have anonymous preferences, and all matchings are individually rational. Further, assume that all match utilities are supermodular. Then for any $(\mu, \mathbf{w}) \in \Gamma^*$, μ is ex ante efficient in $\Gamma^*(\mu)$.*

It is easy to show with examples that ex ante symmetry and supermodularity of match utilities are essential for the previous proposition. If these two assumptions are dropped then one obtains a weaker conclusion:

Proposition 8. *Assume that utility functions are increasing, agents have anonymous preferences, and all matchings are individually rational. Then for any $(\mu, \mathbf{w}) \in \Gamma^*$, μ is interim efficient in $\Gamma^*(\mu)$.*

4.2. Single-crossing preferences

Next, I consider a class of non-anonymous preferences that includes linear utility functions for workers and any increasing utility function for firms.

Worker i 's utility function satisfies *single crossing* if for any firms j and \hat{j} , $u_i(w_i, f_j, j) - u_i(w_i, f_{\hat{j}}, \hat{j})$, $j \neq \hat{j}$, crosses zero at most once as w_i increases from \underline{w} to \bar{w} .¹⁸

Consider a scenario where firm j is a better fit than firm \hat{j} for high-ability workers. Then, one would expect the utility difference $u_i(w, f_j, j) - u_i(w, f_{\hat{j}}, \hat{j})$ to increase with w ; if $u_i(\underline{w}, f_j, j) - u_i(\underline{w}, f_{\hat{j}}, \hat{j}) < 0$ and $u_i(\bar{w}, f_j, j) - u_i(\bar{w}, f_{\hat{j}}, \hat{j}) > 0$ then the utility difference crosses zero exactly once.

Note that the single-crossing assumption does not require that $u_i(w, f_j, j) - u_i(w, f_{\hat{j}}, \hat{j})$ is increasing in w , only that it crosses zero at most once. To simplify the proofs, it is assumed that $u_i(w, f_j, j) - u_i(w, f_{\hat{j}}, \hat{j})$ equals zero for at most one value of w ; i.e., $u_i(w, f_j, j) - u_i(w, f_{\hat{j}}, \hat{j})$ is increasing in w when the utility difference is zero.

Next, a sufficient condition for interim efficiency of a stable matching outcome is obtained.

Proposition 9. *Assume that workers are ex ante symmetric and have single-crossing preferences and that firms' utility functions are increasing. Let B be a Bayesian self-stabilizing set of matching outcomes. Then for any $(\mu, \mathbf{w}) \in B$, μ is interim efficient in $B(\mu)$.¹⁹*

In [Example 3](#) below, the hypotheses of [Proposition 7](#) is satisfied except that the assumption of anonymous preferences is relaxed to single crossing (for workers). There is a μ in Γ^* in this example which is not ex ante efficient in $\Gamma^*(\mu)$. Thus, [Proposition 7](#) cannot be generalized to single-crossing preferences.

¹⁸ If either $j = 0$ or $\hat{j} = 0$ then single crossing follows from increasing utility.

¹⁹ Following the convention for $\Gamma^*(\mu)$, the set $B(\mu) \equiv \{ \mathbf{w} \mid (\mu, \mathbf{w}) \in B \}$.

Example 3. There are two firms and two workers. The workers’ types are i.i.d. uniform on the interval $[0, 1]$. The utility functions of workers are given in the table below:

Worker utility	Firm 1	Firm 2
$u_1(w_1, j)$	$w_1 + 1$	$2w_1$
$u_2(w_2, j)$	$w_2 + 1$	$2w_2$

Each worker prefers firm 1 to firm 2. Firms’ utility functions are $v_j(w_i, i) = w_i$. Thus, agents’ preferences are ex ante symmetric. Firms’ preferences are anonymous while worker preferences satisfy single crossing. Match utilities are weakly supermodular.

Consider matching μ_1 where $\mu_1(1) = 1$ and $\mu_1(2) = 2$. It may be verified that μ_1 is in Γ^* and that

$$\Gamma^*(\mu_1) = \{(w_1, w_2) \mid 0.5 \leq w_1 \leq 1 \text{ and } w_2 \leq 1\}$$

Let μ_2 be the matching where $\mu_2(1) = 2$ and $\mu_2(2) = 1$. Conditional on the set $\Gamma^*(\mu_1)$, the expected surplus under μ_1 minus the expected surplus under μ_2 is

$$2 \int_{0.5}^1 \int_0^1 (w_2 - w_1) dw_2 dw_1 = -0.25$$

Thus, μ_1 is ex ante dominated by μ_2 in $\Gamma^*(\mu_1)$. A similar argument shows that μ_2 is ex ante dominated by μ_1 in $\Gamma^*(\mu_2)$. \square

4.3. Existence and extensions

As Proposition 3 shows, the set of Bayesian stable matching-outcomes, Γ^* , is non-empty. However, we know from Example 2 that there may exist \mathbf{w} such that for any matching μ , $(\mu, \mathbf{w}) \notin \Gamma^*$, i.e., Γ^* may be “locally” empty at some \mathbf{w} . Lemma 3 provides a characterization of \mathbf{w} at which Γ^* is locally nonempty when preferences are anonymous. In the statement of the lemma, w_j^* is as defined in Proposition 6 and $w_{(j)}$ is the j th highest worker type at $\mathbf{w} = (w_1, w_2, \dots, w_n)$.

Lemma 3. Assume that utility functions are increasing, agents have anonymous preferences, and that all matchings are individually rational. Let \mathbf{w} be a realization of worker types. Then there exists a matching μ such that $(\mu, \mathbf{w}) \in \Gamma^*$ if and only if $w_{(j)} \geq w_j^*$, for $j = 1, 2, \dots, \min[m, n]$.

The proof of Lemma 3 points to the difficulty in obtaining existence of stable matchings for all \mathbf{w} when preferences are anonymous. Here, existence fails when worker types are small. Even with more general preferences, existence usually fails if \mathbf{w} is close to $\underline{\mathbf{w}}$: if each firm is matched to a worker with type close to \underline{w} , a Bayesian-block may become attractive to at least one firm. However, the possibility of existence of Bayesian stable matching-outcomes at each \mathbf{w} is demonstrated below in an example with two firms and workers and a divergence in workers’ preferences.

Example 4. There are two firms and two workers. Each firm has anonymous preferences. Each worker has single-crossing preferences. Worker 1 strictly prefers firm 1 to firm 2 if and only if $w_1 \leq \hat{w}_1$ while worker 2 strictly prefers firm 2 to firm 1 if and only if $w_2 \leq \hat{w}_2$. All matchings are individually rational.

Let μ_1 be a matching where $\mu_1(1) = 1$ and $\mu_1(2) = 2$ and μ_2 be another matching where $\mu_2(1) = 2$ and $\mu_2(2) = 1$. It may be verified that $(\mu_1, (w_1, w_2))$ is a Bayesian stable matching-outcome if $w_1 \leq \hat{w}_1$ and $w_2 \leq \hat{w}_2$. Next, suppose that

$$v_1(\hat{w}_2, f_1) \geq E[v_1(w_1, f_1) | w_1 \leq \hat{w}_1] \quad \text{and} \quad v_2(\hat{w}_1, f_2) \geq E[v_2(w_2, f_2) | w_2 \leq \hat{w}_2]$$

If $|\hat{w}_1 - \hat{w}_2|$ is small then these two inequalities are satisfied. Then $(\mu_2, (w_1, w_2))$ is a Bayesian stable matching-outcome if either $w_1 > \hat{w}_1$ or $w_2 > \hat{w}_2$ or both. \square

In general, however, a Bayesian stable matching-outcome may not exist at each \mathbf{w} . Identifying constrained-efficient mechanisms that ultimately lead to a stable matching-outcome at every \mathbf{w} is a promising topic for future work. A path to a stable matching-outcome, stable in a modified sense, when the initial matching is unstable, is sketched out below.

For simplicity, suppose that all matchings are individually rational. Then, starting with a Bayesian unstable matching-outcome (μ_0, \mathbf{w}) a market reaches a matching outcome (μ_r, \mathbf{w}) through a sequence of Bayesian blocks in which agents cannot block with any of their erst-while partners.²⁰ Thus, the path of matching outcomes $(\mu_0, \mathbf{w}), (\mu_1, \mathbf{w}), \dots, (\mu_r, \mathbf{w})$ is such that $\mu_0, \mu_1, \dots, \mu_r$ are distinct. The matching outcome (μ_ℓ, \mathbf{w}) is reached from $(\mu_{\ell-1}, \mathbf{w})$ after exactly one worker-firm pair, (i_ℓ, j_ℓ) , Bayesian blocks $(\mu_{\ell-1}, \mathbf{w})$; firm $\mu_{\ell-1}(i_\ell)$ and worker $v_{\ell-1}(j_\ell)$ are left unmatched at μ_ℓ .²¹ Define history

$$h_r = ((i_1, \mu_0(i_1)), (v_0(j_1), j_1), \dots, ((i_r, \mu_{r-1}(i_r)), (v_{r-1}(j_r), j_r)))$$

as the sequence of worker-firm pairs that were matched to each other at some prior matching on the path to μ_r . A worker-firm pair (i, j) is *admissible at h_r* if $(i, j) \notin h_r$; that is, worker i and firm j are not matched to each other at any $\mu_\ell, \ell = 0, 1, \dots, r - 1$. Then, $(\mu_r, \mathbf{w}) \in A_r$ is Bayesian A_r -blocked conditional on history h_r if there is an admissible worker-firm pair (i, j) satisfying (10) and (11). There is a finite number of matchings and μ_0, \dots, μ_r are distinct. Therefore, for each initial matching-outcome (μ_0, \mathbf{w}) any sequence of matching outcomes obtained through Bayesian blocks must end at a matching outcome to which there are no admissible Bayesian blocking pairs; this final matching outcome is *Bayesian stable conditional on history*. While there exists a conditional Bayesian stable matching at each \mathbf{w} , the set of conditional Bayesian stable matching-outcomes is path-dependent. In particular, the order in which blocks are entertained has a bearing on the efficiency of the final outcome.

A complementary approach to obtaining Bayesian stable matching-outcomes is to allow firms to return to previously-matched workers, at some cost. If worker-firm pair (i, j) blocks matching μ , then firm j would like to go back to worker $v(j)$ if it turns out that $v_j(w_i, f_j, j) < v_j(w_{v(j)}, f_j, v(j))$. Suppose that firm j may return to worker $v(j)$ after paying a cost $c \geq 0$. The cost is incurred by firm j in order to repair its relationship with worker $v(j)$ after j 's temporary abandonment of $v(j)$. The original definition of Bayesian blocking is modified by replacing (11) with

$$E \left[\max[v_j(w'_i, f_j, i), v_j(w_{v(j)}, f_j, v(j)) - c] \mid \mathbf{w}' \in A^{ij}(\mu), w'_{v(j)} = w_{v(j)} \right] > v_j(w_{v(j)}, f_j, v(j)) \tag{12}$$

²⁰ Recall that an assumption underlying Bayesian blocking is that a firm cannot go back to a worker it was previously matched with if it turns out that the firm is ex post worse off in the block.

²¹ As all matchings are individually rational, $i_\ell \neq 0$ and $j_\ell \neq 0$.

A matching-outcome $(\mu, \mathbf{w}) \in A$ is *Bayesian A-blocked with costly return* if (i) there is a pair (i, j) satisfying (10) and (12) and (ii) $v_j(w_i, f_j, i) \geq v_j(w_{v(j)}, f_j, v(j)) - c$. If (ii) does not hold then firm j , after learning that it is worse off in the block, will incur a cost c and return to $v(j)$. In effect, firm j explores the profitability of a block when (i) is satisfied; the block is sustained if (ii) is satisfied.²²

As c decreases, Bayesian blocks are more likely to be explored but less likely to be sustained. Consequently, firms become better informed as they learn the types of workers with whom they explore blocks. In the limit as the cost c goes to zero, a worker-firm pair (i, j) explores a block to a matching μ if²³

$$\Pr \left[v_j(w'_i, f_j, i) > v_j(w_{v(j)}, f_j, v(j)) \mid \mathbf{w}' \in A^{ij}(\mu), w'_{v(j)} = w_{v(j)} \right] > 0$$

Consequently, at any \mathbf{w} a block to a complete-information stable matching at \mathbf{w} might be explored but is never sustained as the cost of return is zero. Hence, the set of matching-outcomes that are Bayesian stable with costly return is locally non-empty in the limit as c goes to zero.

CORRELATED TYPES

In the foregoing analysis, Bayesian stability is investigated under the assumption that worker types are independently and identically distributed. If, instead, worker types are correlated then the definitions of Bayesian blocking and of Γ^* are unchanged. To see this, note that as the joint distribution over worker types is common knowledge, the set of blocked matching-outcomes is common knowledge. As before, the joint distribution is updated by successively eliminating from the support worker-type vectors that are in blocked matching-outcomes; at each stage, the updated joint distribution remains common knowledge. It is straightforward to show that with correlated worker types, Proposition 6(i) holds. However, conditional on any matching outcome (μ, \mathbf{w}) , the thresholds w_j^* in Proposition 6(ii) are functions of $w_{v(j)}$. If the expected (unconditional) w_j^* are increasing in firm types, then firms with higher types will be matched to stochastically larger worker types. Finding reasonable conditions on the joint distribution of worker types under which w_j^* are monotone is a topic for subsequent work.²⁴

5. An ex post incentive-compatible mechanism

A centralized incentive-compatible mechanism is presented in this section. In a private-values model, the deferred-acceptance algorithm with workers proposing ensures that truthful revelation of types is a dominant strategy for workers. However, in the interdependent-values setting of this paper the deferred-acceptance algorithm cannot be directly implemented as firms do not know their preferences over workers. Nevertheless, properties of the

²² In the definition of Bayesian stability analyzed in this paper, the cost c is high enough that (ii) is always satisfied.

²³ This is similar to Lazarova and Dimitrov (2013), in that an agent is willing to block if there is positive probability that he will do better.

²⁴ With i.i.d. types, Bayesian stable matching-outcome sets tend to be locally empty at low values of \mathbf{w} . If worker types are affiliated, a firm matched to a worker with a low type is likely to become pessimistic about the types of all workers, making potential blocks less attractive than if worker types were independent. This may reduce the incidence of local emptiness at low types.

set of complete-information stable matchings can be exploited to construct an efficient, ex post incentive-compatible mechanism. For almost all worker types \mathbf{w} , this mechanism implements the worker-optimal complete-information stable matching. However, it is required that the mechanism designer knows agents' utility functions, which is a strong assumption.

The rules of the *worker-optimal mechanism* are as follows:

1. Workers report their types to the mechanism designer.
2. The mechanism designer computes the worker-optimal complete-information stable matching (if one exists) for the reported worker types, pairs the workers and firms according to this stable matching, and reveals the reported worker types to all firms.²⁵
3. After the matching is implemented, each firm reports the true type of its matched worker, if it is different from the reported type.²⁶
4. If the reports made by workers about their types in step 2 coincide with the worker types reported by the firms in step 3, then proceed to step 5. Otherwise, the worker-optimal complete-information stable matching for the types reported by firms is computed and implemented.
5. End of mechanism.

Proposition 10. *Assume that utility is increasing. For almost all worker types, it is an ex post equilibrium for workers to truthfully report their types and for firms to truthfully report the types of their matched workers in the worker-optimal mechanism.*

6. Concluding remarks

Of the two notions of stability under one-sided incomplete information investigated in this paper, ex ante stability imposes a minimal restriction on matching outcomes. Ex ante stable matching-outcomes may not be close to complete-information stable matching-outcomes when the asymmetry of information is small. Consequently, ex ante stability is not a satisfactory desideratum for NTU matching models with incomplete information. Bayesian stability is a more selective benchmark and matching outcomes that satisfy this criterion have sound efficiency properties. The price of this selectivity is that there exist worker types at which there is no Bayesian stable matching. Bayesian blocking with costly return may reduce the incidence of non-existence.

The resting point of a matching market at worker-type vectors at which there is no Bayesian stable matching-outcome is not addressed in this paper. The question as to how a Bayesian stable matching-outcome, at worker types at which it exists, is established is also outside the scope of this paper. The investigation of decentralized mechanisms that address these questions is a next step.

²⁵ A worker-optimal stable matching exists under strict preferences. If workers are truthful, then for almost all worker types, firms and workers have strict preferences as utility is increasing in worker type. If preferences are not strict at the reported types and a worker-optimal matching does not exist, then the mechanism designer picks an arbitrary complete-information stable matching.

²⁶ As before, after a matching is implemented each firm will learn its matched worker's type and will find out whether it is different from the worker type revealed by the mechanism designer.

Acknowledgments

I am grateful to Oleg Baranov, Pradeep Dubey, Federico Echenique, Vijay Krishna, Pino Lopomo, George Mailath, Herve Moulin, Joe Ostroy, Marek Pycia, Eric Shi, Tayfun Sonmez, Utku Unver, three anonymous referees, and an associate editor for helpful comments and suggestions.

Appendix A. Proofs

Proof of Lemma 1. (i) It follows from (3) that $B^{ij}(\mu) \subseteq A^{ij}(\mu)$. Therefore, if $(\mu, \mathbf{w}) \in B \subset A$ is A -blocked for all admissible types then (μ, \mathbf{w}) is also B -blocked for all admissible types.

(ii) Let $B = \{(\mu_{\mathbf{w}}, \mathbf{w})\}$. As $\mu_{\mathbf{w}}$ is complete-information stable at \mathbf{w} , $(\mu_{\mathbf{w}}, \mathbf{w})$ is ex ante B -stable. Therefore, by (i), $(\mu_{\mathbf{w}}, \mathbf{w})$ is ex ante A -stable for any $A \supset B$. \square

Proof of Lemma 2. Suppose A is ex ante self-stabilizing. Therefore, each matching outcome in A is individually rational, i.e., $A \subseteq \Sigma^0$. Let $k - 1 \geq 0$ be such that $A \subseteq \Sigma^{k-1}$. As every $(\mu, \mathbf{w}) \in A$ is ex ante A -stable, it is also ex ante Σ^{k-1} -stable by Lemma 1(i). Thus, $A \subseteq \Sigma^k$. We have $A \subseteq \Sigma^k, \forall k$ by induction. \square

Proof of Proposition 1. (i) Let μ be a maximal matching at which firms of the highest types are matched. By assumption, all matchings are individually rational. Thus, $(\mu, \mathbf{w}) \in \Sigma^0$ for all \mathbf{w} and, in particular, for all $\mathbf{w} > \underline{\mathbf{w}}$. The induction hypothesis is that for some $k \geq 0$, we have $(\mu, \mathbf{w}) \in \Sigma^k$ for all $\mathbf{w} > \underline{\mathbf{w}}$.

Suppose that (4) is satisfied by a pair (i, j) for some $(\mu, \mathbf{w}) \in \Sigma^k, \mathbf{w} > \underline{\mathbf{w}}$. Anonymous preferences and increasing utility imply that $f_j > f_{\mu(i)}$ and therefore $\mathbf{w}' \in A^{ij}(\mu)$ for all \mathbf{w}' such that $w'_i > \underline{w}$. Suppose that $\mu(i) \neq 0$. Hence, $f_j > f_{\mu(i)}$ implies that $v(j) \neq 0$ as firms of the highest types are matched at μ . As $\mathbf{w} > \underline{\mathbf{w}}$ we have $w_{v(j)} > \underline{w}$. Select $w'_i \in (\underline{w}, w_{v(j)})$ and any $w_{-i-v(j)} > (\underline{w}, \dots, \underline{w})$. As $\mathbf{w}' \equiv (w'_i, w_{v(j)}, w_{-i-v(j)}) > \underline{\mathbf{w}}$, under the induction hypothesis we have $(\mu, \mathbf{w}') \in \Sigma^k$. Moreover, $\mathbf{w}' \in A^{ij}(\mu)$ and $w_{v(j)} = w'_{v(j)}$ but, by increasing utility and anonymous preferences, $v_j(w_{v(j)}, f_j) > v_j(w'_i, f_j)$. Thus, (5) is not satisfied. Hence, (μ, \mathbf{w}) is not Σ^k -blocked for all admissible types.

If, instead, (i, j) satisfies (4) at $(\mu, \mathbf{w}) \in \Sigma^k$ and $\mu(i) = 0$ then, as μ is maximal, all firms are matched at μ . Therefore, $v(j) \neq 0$ and the argument in the previous paragraph implies that (μ, \mathbf{w}) is not Σ^k -blocked for all admissible types.

If, instead, (4) is not satisfied by any pair (i, j) for any $(\mu, \mathbf{w}) \in \Sigma^k, \mathbf{w} > \underline{\mathbf{w}}$, then again (μ, \mathbf{w}) is not Σ^k -blocked for all admissible types.

Thus, $(\mu, \mathbf{w}) \in \Sigma^{k+1}$. By induction, $(\mu, \mathbf{w}) \in \Sigma^k$, for all k , for all $\mathbf{w} > \underline{\mathbf{w}}$. Consequently, $(\mu, \mathbf{w}) \in \Sigma^*$, for all $\mathbf{w} > \underline{\mathbf{w}}$.

(ii) As $\mathbf{w} > \underline{\mathbf{w}}$, all matchings are individually rational, and utility is increasing, every agent strictly prefers to be matched rather than remain unmatched. If matching μ is not maximal then there exists an unmatched worker and an unmatched firm; this worker-firm pair Σ^0 -blocks (μ, \mathbf{w}) .

If firms of the highest types are not matched at μ , then there exist firms j and \hat{j} with $f_{\hat{j}} > f_j$ such that firm \hat{j} is unmatched and firm j is matched at μ . Then $(v(j), \hat{j}) \Sigma^0$ -blocks (μ, \mathbf{w}) . \square

Proof of Proposition 3. Let $\bar{\mu}$ be a complete-information stable matching at $\bar{\mathbf{w}} = (\bar{w}, \bar{w}, \dots, \bar{w})$. It is shown that $(\bar{\mu}, \bar{\mathbf{w}}) \in \Gamma^k$ for all k .

By individually rationality, $(\bar{\mu}, \bar{\mathbf{w}}) \in \Gamma^0$. Suppose that $(\bar{\mu}, \bar{\mathbf{w}}) \in \Gamma^k$ is Bayesian Γ^k -blocked by a worker-firm pair (i, j) . Therefore,

$$\bar{\mathbf{w}} \in \Gamma^{kij}(\mu), \quad \text{i.e.,} \quad u_i(\bar{w}_i, f_j, j) > u_i(\bar{w}_i, f_{\mu(i)}, \mu(i))$$

$$\text{and } E[v_j(w'_i, f_j, i) | \mathbf{w}' \in \Gamma^{kij}(\mu), w'_{\mu(j)} = w_{\mu(j)}] > v_j(\bar{w}_{v(j)}, f_j, v(j))$$

where $\Gamma^{kij}(\mu)$ is defined using (3). By increasing utility,

$$v_j(\bar{w}_i, f_j, i) \geq E[v_j(w'_i, f_j, i) | \mathbf{w}' \in \Gamma^{kij}(\mu), w'_{\mu(j)} = w_{\mu(j)}] > v_j(\bar{w}_{v(j)}, f_j, v(j))$$

Thus, $\bar{\mu}$ is complete-information blocked by worker i and firm j , contradicting the assumption that $\bar{\mu}$ is a complete-information stable matching at $\bar{\mathbf{w}}$.

Therefore, $(\bar{\mu}, \bar{\mathbf{w}}) \in \Gamma^{k+1}$ and thus, $(\bar{\mu}, \bar{\mathbf{w}}) \in \Gamma^*$. \square

Proof of Proposition 5. As firms' utility functions are increasing, by Propositions 2 and 4 only two-person blocking coalitions need to be considered.

Suppose that $(\mu, \mathbf{w}) \in B$ is not Bayesian blocked in B . Thus, for any worker-firm pair (i, j) satisfying $u_i(w_i, f_j, j) > u_i(w_i, f_{\mu(i)}, \mu(i))$ we have

$$E[v_j(w'_i, f_j, i) | \mathbf{w}' \in B^{ij}(\mu), w'_{\mu(j)} = w_{\mu(j)}] \leq v_j(w_{v(j)}, f_j, v(j))$$

Hence, there exists $(w'_i, w'_{-i}) \in B^{ij}(\mu)$ such that $v_j(w'_i, f_j, i) \leq v_j(w_{v(j)}, f_j, v(j))$. Thus, (μ, \mathbf{w}) is not B -blocked for all admissible types and, by Lemma 1(i), (μ, \mathbf{w}) is not A -blocked for all admissible types.

As $\Gamma^0 = \Sigma^0$, we have $\Gamma^k \subseteq \Sigma^k$, for all k and hence $\Gamma^* \subseteq \Sigma^*$. \square

Proof of Proposition 6. (i) If μ is not maximal then any (μ, \mathbf{w}) is Bayesian blocked by an unmatched worker-firm pair. If μ is such that a higher-type firm is unmatched then any (μ, \mathbf{w}) is Bayesian blocked by an unmatched higher-type firm and a worker matched to a lower-type firm. Conversely, if μ is a maximal matching in which firms of the highest types are matched then it is easily verified that $(\mu, \bar{\mathbf{w}}) \in \Gamma^*$.

(ii) Assume that $f_{j-1} > f_j$ for all j . A proof for the case $f_{j-1} = f_j$ for some j follows easily.

EQUAL NUMBERS OF WORKS AND FIRMS: $m = n$. Let μ be a maximal matching.²⁷ First, note that

$$\Gamma^0(\mu) \equiv \{\mathbf{w} \mid w_{v(j)} \geq \underline{w}, \forall j\}.$$

The induction hypothesis is that there exist $w_{v(j)}^{k-1} < \bar{w}$ for all j such that $\Gamma^{k-1}(\mu) = \{\mathbf{w} \mid w_{v(j)} \geq w_{v(j)}^{k-1}, \forall j\}$. This is satisfied for $k - 1 = 0$, with $w_{v(j)}^0 = \underline{w}$ for all j .

No worker is willing to block with firm n because $f_n < f_j$ for all $j < n$. As firm n can never be part of a blocking pair, the fact that (μ, \mathbf{w}) is unblocked in $\Gamma^0, \Gamma^1, \dots, \Gamma^{k-2}$, i.e., $\mathbf{w} \in \Gamma^{k-1}(\mu)$, conveys no information about $w_{v(n)}$. Hence, $w_{v(n)}^{k-1} = \underline{w}$.

For $j < n$ and any $\ell > j$, worker $v(\ell)$ is willing to participate in a block with firm j as $f_j > f_\ell$. Let $\ell^m = \arg \max_{\ell > j} w_{v(\ell)}^{k-1}$. Define

²⁷ Non-maximal matchings are also individually rational. For simplicity, such matchings are excluded from consideration as they are blocked by any unmatched worker-firm pair.

$$w_{v(j)}^k \equiv \inf \left\{ w \in [\underline{w}, \bar{w}] \mid v_j(w, f_j) \geq \mathbb{E} \left[v_j(w, f_j) \mid w \geq w_{v(\ell^m)}^{k-1} \right] \right\}$$

As v_j is increasing in w and $w_{v(\ell^m)}^{k-1} < \bar{w}$ by the induction hypothesis, there exists such a $w_{v(j)}^k \in (w_{v(\ell^m)}^{k-1}, \bar{w})$.

If $w_{v(j)} < w_{v(j)}^k$ then firm j and worker $v(\ell^m)$ form a blocking pair to μ . If, instead, $w_{v(j)} \geq w_{v(j)}^k$ then firm j is not willing to block with any worker $v(\ell)$, $\ell > j$. No worker $v(\ell)$ with $\ell < j$ would block with firm j as $f_j < f_\ell$. Consequently, if firm j does not participate in a block at stage k , it is clear that $w_{v(j)} \geq w_{v(j)}^k$. Thus, $\Gamma^k(\mu) = \{\mathbf{w} \mid w_{v(j)} \geq w_{v(j)}^k, \forall j\}$.

By induction, we have proved that for all k and j there exist $w_{v(j)}^k < \bar{w}$ such that $\Gamma^k(\mu) = \{\mathbf{w} \mid w_{v(j)} \geq w_{v(j)}^k, \forall j\}$. Further, by definition $w_{v(j)}^k \geq w_{v(j)}^{k-1}$. As preferences are anonymous, $w_{v(j)}^k$ does not depend on the identity of worker $v(j)$. Hence, $w_j^* \equiv \lim_{k \uparrow \infty} w_{v(j)}^k$ is well-defined and we have

$$\Gamma^*(\mu) = \{\mathbf{w} \mid w_{v(j)} \geq w_j^*\}.$$

Next, suppose that $w_{j-1}^* \leq w_j^*$ for some j . Then for small enough $\epsilon > 0$,

$$\begin{aligned} \mathbb{E} \left[v_{j-1}(w_{v(j)}, f_{j-1}) \mid w_{v(j)} \geq w_j^* \right] &> v_{j-1}(w_{v(j-1)}, f_{j-1}), \\ \forall w_{v(j-1)} \in [w_{j-1}^*, w_{j-1}^* + \epsilon]. \end{aligned}$$

Thus, as $f_{j-1} > f_j$, $(v(j), j - 1)$ form a blocking pair if $w_{v(j-1)} \in [w_{j-1}^*, w_{j-1}^* + \epsilon]$. But this contradicts the fact that Γ^* is a self-stabilizing set. Hence, $w_{j-1}^* > w_j^*$ for all j . With $w_n^* = \underline{w}$, the remaining w_j^* , $j < n$ are defined recursively by

$$w_j^* \equiv \inf \left\{ w \in [\underline{w}, \bar{w}] \mid v_j(w, f_j) \geq \mathbb{E} \left[v_j(w_{v(j+1)}, f_j) \mid w_{v(j+1)} \geq w_{j+1}^* \right] \right\} \tag{13}$$

MORE FIRMS THAN WORKERS: $m > n$. The $m - n$ firms with the lowest types $f_{n+1}, f_{n+2}, \dots, f_m$ are not matched at any $(\mu, \mathbf{w}) \in \Gamma^*$. For $j = 1, \dots, n$, firm j is matched to worker $v(j)$ with $w_{v(j)} \geq w_j^*$, where w_j^* , $j < n$ are defined recursively by (13), with $w_n^* = \underline{w}$.

MORE WORKERS THAN FIRMS: $m < n$. All firms are matched at any $(\mu, \mathbf{w}) \in \Gamma^*$, with w_n^* satisfying $\mathbb{E}[v_n(w, f_n)] = v_n(w_n^*, f_n)$ and w_j^* , $j < n$ defined recursively by (13). \square

Proof of Proposition 7. Under the assumptions of ex ante symmetry and anonymous preferences, a matching μ is ex ante efficient in $\Gamma^*(\mu)$ if for any other matching $\hat{\mu}$,

$$\begin{aligned} &\mathbb{E} \left[\sum_{i=1}^n u(w_i, f_{\mu(i)}) + \sum_{j=1}^m v(w_{v(j)}, f_j) \mid \mathbf{w} \in \Gamma^*(\mu) \right] \\ &\geq \mathbb{E} \left[\sum_{i=1}^n u(w_i, f_{\hat{\mu}(i)}) + \sum_{j=1}^m v(w_{\hat{v}(i)}, f_j) \mid \mathbf{w} \in \Gamma^*(\mu) \right] \end{aligned} \tag{14}$$

Assume, without loss of generality, that $f_1 \geq f_2 \geq \dots \geq f_m$.

EQUAL NUMBERS OF WORKS AND FIRMS: $m = n$. Consider any (μ, \mathbf{w}) in Γ^* . By Proposition 6, all workers and firms are matched at μ and there exist $w_{j-1}^* \geq w_j^*$ such that $\Gamma^*(\mu) =$

$\{\mathbf{w} \mid w_{v(j)} \geq w_j^*, \forall j\}$. It will be convenient to use notation $w_{v(j)}^*$ instead of w_j^* . Thus, $\Gamma^*(\mu) = \{\mathbf{w} \mid w_{v(j)} \geq w_{v(j)}^*, \forall j\}$ and (14) may be written as

$$\begin{aligned} & \sum_{j=1}^m \mathbb{E} \left[u(w_{v(j)}, f_j) + v(w_{v(j)}, f_j) \mid w_{v(j)} \geq w_{v(j)}^* \right] \\ & \geq \sum_{j=1}^m \mathbb{E} \left[u(w_{\hat{v}(j)}, f_j) + v(w_{\hat{v}(j)}, f_j) \mid w_{\hat{v}(j)} \geq w_{v(j)}^* \right] \end{aligned} \tag{15}$$

The following definition is useful. An inverse matching $v_b = (v_b(1), v_b(2), \dots, v_b(n))$ is a *circular permutation* of another inverse matching $v_a = (v_a(1), v_a(2), \dots, v_a(n))$ if there exist k, k' , with $k \leq k'$ such that

$$\begin{aligned} v_b(k) &= v_a(k'), & v_b(\ell) &= v_a(\ell - 1), & \ell &= k + 1, \dots, k' \\ v_b(\ell) &= v_a(\ell), & \forall \ell < k & \text{ and } \forall \ell > k' \end{aligned}$$

That is, v_b is obtained from v_a by replacing the k th element in the vector v_a by its k' th element and shifting down one place the k th through the $(k' - 1)$ th elements in v_a . If $k = k'$ then $v_a = v_b$.

It is sufficient to show that (15) holds for any matching $\hat{\mu}$ at which all firms are matched. At the corresponding inverse matching $\hat{v}, (\hat{v}(1), \hat{v}(2), \dots, \hat{v}(n))$ is a permutation of $(v(1), v(2), \dots, v(n))$.²⁸ We show by construction that \hat{v} can be reached from v through (at most) $n - 1$ circular permutations. That is, there exist $v = v_0, v_1, \dots, v_{n-1} = \hat{v}$ where v_r is a circular permutation of $v_{r-1}, r = 1, 2, \dots, n - 1$. Further, for all r ,

$$v_r(\ell) = \hat{v}(\ell), \quad \forall \ell \leq r, \quad \text{and} \quad w_{v_r(\ell)}^* \geq w_{v_r(\ell+1)}^*, \quad \forall \ell > r. \tag{16}$$

Condition (16) is trivially true for v_0 . Assume that it is true for matchings $v_0, v_1, \dots, v_{r-1}, r < n - 1$. We show that there exists v_r , a circular permutation of v_{r-1} , for which (16) is true. Let r' be the smallest $r' \geq r$ such that $\hat{v}(r) = v_{r-1}(r')$. As v_{r-1} is a permutation of v , and therefore also a permutation of \hat{v} , r' exists. Define,

$$v_r(\ell) = \begin{cases} v_{r-1}(\ell), & \text{if } \ell < r \text{ or } \ell > r' \\ v_{r-1}(r'), & \text{if } \ell = r \\ v_{r-1}(\ell - 1), & \text{if } r < \ell \leq r' \end{cases}$$

The matching v_r is obtained from v_{r-1} by replacing the element $v_{r-1}(r)$ (i.e., the worker matched to the firm with type f_r in the matching v_{r-1}) with $v_{r-1}(r')$ and sliding down one place each element $v_{r-1}(\ell), \ell = r, \dots, r' - 1$. Thus, v_r is a circular permutation of v_{r-1} . It is easily verified that v_r satisfies (16). Proceeding in this fashion, we have $v_{n-1} = \hat{v}$.

To complete the proof (for the case $m = n$), we show that the expected surplus in $\Gamma^*(\mu)$ under v_{r-1} is at least as large as under v_r . If $v_r = v_{r-1}$ then there is nothing to prove. Therefore, suppose that $v_r \neq v_{r-1}$, i.e., $r' > r$. Let $M(w, f) \equiv u(w, f) + v(w, f)$ be the match utility when a worker of type w and a firm of type f are matched. Then, the difference between the expected welfare under v_{r-1} and under v_r is:

²⁸ In the rest of the proof, inverse matchings v and \hat{v} are used instead of μ and $\hat{\mu}$.

$$\begin{aligned}
 & \sum_{j=1}^m \mathbb{E} \left[M(w_{v_{r-1}(j)}, f_j) \mid w_{v_{r-1}(j)} \geq w_{v_{r-1}(j)}^* \right] - \sum_{j=1}^m \mathbb{E} \left[M(w_{v_r(j)}, f_j) \mid w_{v_r(j)} \geq w_{v_r(j)}^* \right] \\
 &= \sum_{j=r}^{r'} \mathbb{E} \left[M(w_{v_{r-1}(j)}, f_j) \mid w_{v_{r-1}(j)} \geq w_{v_{r-1}(j)}^* \right] - \sum_{j=r}^{r'} \mathbb{E} \left[M(w_{v_r(j)}, f_j) \mid w_{v_r(j)} \geq w_{v_r(j)}^* \right] \\
 &= \sum_{j=r}^{r'-1} \mathbb{E} \left[M(w_{v_{r-1}(j)}, f_j) - M(w_{v_{r-1}(j)}, f_{j+1}) \mid w_{v_{r-1}(j)} \geq w_{v_{r-1}(j)}^* \right] \\
 &\quad + \mathbb{E} \left[M(w_{v_{r-1}(r')}, f_{r'}) - M(w_{v_{r-1}(r')}, f_r) \mid w_{v_{r-1}(r')} \geq w_{v_{r-1}(r')}^* \right] \\
 &\geq \sum_{j=r}^{r'-1} \mathbb{E} \left[M(w_{v_{r-1}(j)}, f_j) - M(w_{v_{r-1}(j)}, f_{j+1}) \mid w_{v_{r-1}(j)} \geq w_{v_{r-1}(r')}^* \right] \\
 &\quad + \mathbb{E} \left[M(w_{v_{r-1}(r')}, f_{r'}) - M(w_{v_{r-1}(r')}, f_r) \mid w_{v_{r-1}(r')} \geq w_{v_{r-1}(r')}^* \right] \\
 &= \sum_{j=r}^{r'-1} \mathbb{E} \left[M(w, f_j) - M(w, f_{j+1}) \mid w \geq w_{v_{r-1}(r')}^* \right] \\
 &\quad + \mathbb{E} \left[M(w, f_{r'}) - M(w, f_r) \mid w \geq w_{v_{r-1}(r')}^* \right] \\
 &= 0
 \end{aligned}$$

The supermodularity of $M(w, f)$ and $f_j \geq f_{j+1}$ implies that $M(w_{v_{r-1}(j)}, f_j) - M(w_{v_{r-1}(j)}, f_{j+1})$ is an increasing function of $w_{v_{r-1}(j)}$. Thus, the inequality follows because $w_{v_{r-1}(j)}^* \geq w_{v_{r-1}(r')}^*$, $j = r, \dots, r' - 1$, implies that the probability distribution over $w_{v_{r-1}(j)}$ conditional on $w_{v_{r-1}(j)} \geq w_{v_{r-1}(j)}^*$ dominates by first-order stochastic dominance the probability distribution over $w_{v_{r-1}(j)}$ conditional on $w_{v_{r-1}(j)} \geq w_{v_{r-1}(r')}^*$.

MORE FIRMS THAN WORKERS: $m > n$. At any μ in Γ^* , μ is maximal and firms with $m - n$ lowest types are not matched. The preceding proof establishes that μ generates greater expected welfare in $\Gamma^*(\mu)$ than any $\hat{\mu}$ at which firms with $m - n$ lowest types are not matched. Further, at any $\hat{\mu}$ at which one or more of the firms with $m - n$ lowest types are matched is ex ante dominated by a matching in which none of the $m - n$ lowest type firms are matched.

MORE WORKERS THAN FIRMS: $m < n$. The proof for $m = n$ establishes that μ generates greater expected welfare in $\Gamma^*(\mu)$ than any $\hat{\mu}$ at which the same $n - m$ workers are not matched. Moreover, as the marginal distribution over the type of a worker matched at μ first-order stochastically dominates the type of a worker not matched at μ , μ also generates greater expected welfare in $\Gamma^*(\mu)$ than any $\hat{\mu}$ at which the set of workers not matched differs from the corresponding set at μ . \square

Proof of Proposition 8. Under anonymous preferences, the identity of the matched agent does not enter the utility function. Thus, for any μ in Γ^* , interim efficiency in $\Gamma^*(\mu)$ requires that there not exist another matching $\hat{\mu}$ such that for all $\mathbf{w} \in \Gamma^*(\mu)$

$$u_i(w_i, f_{\hat{\mu}(i)}) \geq u_i(w_i, f_{\mu(i)}), \forall i, \text{ and } E\left[v_j(w_{\hat{v}(j)}, f_j) \mid \mathbf{w} \in \Gamma^*(\mu)\right] \geq v_j(w_{v(j)}, f_j), \forall j$$

with at least one strict inequality.

Let μ be a matching in Γ^* . Suppose that there exists a matching $\hat{\mu}$ and worker i such that $u_i(w_i, f_{\hat{\mu}(i)}) > u_i(w_i, f_{\mu(i)})$ for some $(w_i, w_{-i}) \in \Gamma^*(\mu)$. By increasing utility, $f_{\hat{\mu}(i)} > f_{\mu(i)}$. If $\mu(i) = 0$, i.e., worker i is unmatched at μ , then $\hat{\mu}(i) \neq 0$ and, because μ is maximal by Proposition 6, there exists another worker who is matched at μ but not at $\hat{\mu}$; this worker is strictly worse off at $\hat{\mu}$ than at μ . Consequently, $\hat{\mu}$ cannot interim dominate μ . Hence, assume that $\mu(i) \neq 0$. Assume, without loss of generality, that either $f_{\hat{\mu}(i)} \geq f_j$ for all j or that any firm j with $f_j > f_{\hat{\mu}(i)}$ is matched to the same worker under μ and $\hat{\mu}$. Let worker \hat{i} be such that $\mu(\hat{i}) = \hat{\mu}(i)$. Clearly, $\mu(\hat{i}) \neq \hat{\mu}(\hat{i})$. Further, $f_{\mu(\hat{i})} \geq f_{\hat{\mu}(\hat{i})}$ as any firm j with $f_j > f_{\mu(\hat{i})} = f_{\hat{\mu}(i)}$ is matched to the same worker under μ and $\hat{\mu}$.

Suppose that for all other firms $j \neq \mu(\hat{i})$ we have $f_j \neq f_{\mu(\hat{i})}$. Hence, $f_{\mu(\hat{i})} > f_{\hat{\mu}(\hat{i})}$, and therefore $u_i(w_{\hat{i}}, f_{\hat{\mu}(\hat{i})}) < u_i(w_{\hat{i}}, f_{\mu(\hat{i})})$. Thus, $\hat{\mu}$ does not interim dominate μ .

Suppose, instead, there exists exactly one firm $j \neq \mu(\hat{i})$ such that $f_j = f_{\mu(\hat{i})}$. As μ is in Γ^* and $f_j = f_{\mu(\hat{i})} > f_{\mu(i)}$, there is a worker i' such that $\mu(i') = j$; this follows from the fact that, by Proposition 6, firms of the highest types are matched at μ . Then, the argument in the preceding paragraph implies that either $f_{\mu(\hat{i})} > f_{\hat{\mu}(\hat{i})}$ or $f_{\mu(i')} > f_{\hat{\mu}(i')}$ (or both). Therefore either \hat{i} or i' is worse off at $\hat{\mu}$ than at μ . Once again, $\hat{\mu}$ does not interim dominate μ . Finally, if there is more than one firm $j \neq \mu(\hat{i})$ such that $f_j = f_{\mu(\hat{i})}$, then each of these firms must be matched to a worker at μ and the same argument establishes that at least one of these workers is worse off at $\hat{\mu}$ than at μ .

The only remaining possibility is that μ is interim dominated by a matching $\hat{\mu}$ at which all workers get the same utility as at μ at each $\mathbf{w} \in \Gamma^*(\mu)$ but at least one firm is better off for one $\mathbf{w} \in \Gamma^*(\mu)$. Thus, consider a $\hat{\mu}$ such that for all $(w_i, w_{-i}) \in \Gamma^*(\mu)$, $u_i(w_i, f_{\hat{\mu}(i)}) = u_i(w_i, f_{\mu(i)})$ for all i . As $\hat{\mu} \neq \mu$, there exists a worker i' such that $\hat{\mu}(i') \neq \mu(i')$. Further, as $u_{i'}(w_{i'}, f_{\hat{\mu}(i')}) = u_{i'}(w_{i'}, f_{\mu(i')})$ for all $(w_{i'}, w_{-i'}) \in \Gamma^*(\mu)$, we have $\hat{\mu}(i') \neq 0 \neq \mu(i')$. That is, i' is matched to a firm at both μ and $\hat{\mu}$ and $f_{\hat{\mu}(i')} = f_{\mu(i')}$. Let $j = \mu(i')$. If $\hat{v}(j) = 0$ then firm j is worse at $\hat{\mu}$ than at μ and we are done. Therefore, assume that $\hat{v}(j) \neq 0$. By Proposition 6, we have $(\bar{w} - \epsilon, w_{-i}) \in \Gamma^*(\mu)$ for ϵ arbitrarily small and there exists $w^a < \bar{w}$ such that $(w_{\hat{v}(j)}, w_{-\hat{v}(j)}) \in \Gamma^*(\mu)$ for all $w_{\hat{v}(j)} \geq w^a$. But then, for small enough ϵ , firm $\mu(i')$ is worse off under $\hat{\mu}$ as

$$E\left[v_{\mu(i')}(w_{\hat{v}(j)}, f_{\mu(i')}) \mid w_{\hat{v}(j)} \geq w^a\right] < v_{\mu(i')}(\bar{w} - \epsilon, f_{\mu(i')})$$

Consequently, there does not exist $\hat{\mu}$ that interim dominate μ . \square

Proof of Proposition 9. Consider a matching μ in B , where B is a Bayesian self-stabilizing set of matching outcomes. Let $\hat{\mu}$ be another matching. By single crossing we know that for each worker i such that $\mu(i) \neq \hat{\mu}(i)$, the utility difference²⁹ $u(w_i, f_{\hat{\mu}(i)}, \hat{\mu}(i)) - u(w_i, f_{\mu(i)}, \mu(i))$ crosses zero (at most) once, either from below or from above. Suppose that there exists a worker ℓ whose utility difference between $\hat{\mu}(\ell)$ and $\mu(\ell)$ crosses zero from below. That is, $u(w_\ell, f_{\hat{\mu}(\ell)}, \hat{\mu}(\ell)) > u(w_\ell, f_{\mu(\ell)}, \mu(\ell))$ if and only if $w_\ell > w_\ell^*$ for some $w_\ell^* \in [\underline{w}, \bar{w}]$. Let $\hat{j} = \hat{\mu}(\ell)$. Then, for any $\mathbf{w}' \in B(\mu)$ such that $u(w'_\ell, f_{\hat{j}}, \hat{j}) > u(w'_\ell, f_{\mu(\ell)}, \mu(\ell))$ we have

²⁹ The subscript on worker utility functions is dropped by the assumption of ex ante symmetry of workers.

$$v_{\hat{j}}(w_{v(\hat{j})}, f_{\hat{j}}, v(\hat{j})) \geq E \left[v_{\hat{j}}(w'_{\ell}, f_{\hat{j}}, \ell) \mid \mathbf{w}' \in B, w'_{v(\hat{j})} = w_{v(\hat{j})}, \text{ s.t. } w_{\ell} > w_{\ell}^* \right]$$

$$> E \left[v_{\hat{j}}(w'_{\ell}, f_{\hat{j}}, \ell) \mid \mathbf{w}' \in B, w'_{v(\hat{j})} = w_{v(\hat{j})} \right]$$

where the first inequality follows from the fact that μ is unblocked in B and the second inequality follows from the fact that firm j 's utility function is increasing in w_{ℓ} . Thus, $\hat{\mu}$ does not interim dominate μ in $B(\mu)$.

The rest of the proof establishes that there exists a worker whose utility difference between its matches at $\hat{\mu}$ and μ satisfies single crossing from below.

If there is a firm (worker) that is matched to a worker (firm) in μ but not in $\hat{\mu}$ then $\hat{\mu}$ does not interim dominate μ . Thus, we can restrict attention to the case where $\hat{\mu}$ is a permutation of the non-zero elements of μ . Consequently, there exist workers $i_1, i_2, \dots, i_k, i_{k+1} = i_1$ such that

$$\hat{\mu}(i_p) = \mu(i_{p+1}) \neq 0, \quad p = 1, 2, \dots, k$$

Thus, $u(w, f_{\hat{\mu}(i_p)}, \hat{\mu}(i_p)) - u(w, f_{\mu(i_p)}, \mu(i_p)) = u(w, f_{\mu(i_{p+1})}, \mu(i_{p+1})) - u(w, f_{\mu(i_p)}, \mu(i_p))$ and by single crossing, this utility difference crosses zero once. Suppose that for each p , this zero crossing is from above. That is, for each p there exists $w_{i_p}^* \in (\underline{w}, \bar{w})$ such that $u(w, \mu(i_{p+1})) - u(w, \mu(i_p)) > 0$ if and only if $w < w_{i_p}^*$. Taking $w < \min_p \{w_{i_p}^*\}$, we have a contradiction as

$$0 = \sum_{p=1}^k \left[u(w, f_{\mu(i_{p+1})}, \mu(i_{p+1})) - u(w, f_{\mu(i_p)}, \mu(i_p)) \right] > 0$$

Thus, there exists a worker i_p such that $u(w, f_{\hat{\mu}(i_p)}, \hat{\mu}(i_p)) - u(w, f_{\mu(i_p)}, \mu(i_p))$ crosses zero from below. \square

Proof of Lemma 3. Assume without loss of generality that $f_1 \geq f_2 \geq \dots \geq f_m$. Suppose that $w_{(j)} \geq w_j^*$, for $j = 1, 2, \dots, \min[m, n]$. Let μ^* be any maximal matching at which firm $j \leq n$ is matched to a worker with type greater than or equal to w_j^* . That a μ^* exists follows from $w_{(j)} \geq w_j^*$ for all j ; in particular matching firm j with the worker with the j th highest type is one such matching. If $m > n$ then firms $n + 1, \dots, m$ are unmatched. Then $(\mu, \mathbf{w}) \in \Gamma^*$ as it satisfies Proposition 6(i) & (ii).

Conversely, suppose that there exists $j \leq \min[m, n]$ such that $w_{(j)} < w_j^*$. Hence, at any matching μ which satisfies Proposition 6(i) and $w_{v(\hat{j})} \geq w_{\hat{j}}^*$ for all $\hat{j} < j$, we have $w_{v(j)} < w_j^*$. Therefore, Proposition 6(ii) is not satisfied and (μ, \mathbf{w}) is not a stable matching-outcome. \square

Proof of Proposition 10. Let the realized vector of worker types, $\mathbf{w} = (w_1, w_2, \dots, w_n)$, be such that all agents have strict preference. (As utility is increasing, this is true for almost all type vectors.) Therefore, from Gale and Shapley (1962) we know that a worker-optimal stable matching exists at \mathbf{w} ; call it $\mu_{\mathbf{w}}$.

Suppose that worker 1, say, lies and reports a type $w'_1 \neq w_1$. Let μ' be the matching implemented at $\mathbf{w}' = (w'_1, w_{-1})$.³⁰ If $\mu'(1) = \mu_{\mathbf{w}}(1)$, then worker 1 does not benefit from this deviation. Therefore, suppose that $\mu'(1) \neq \mu_{\mathbf{w}}(1)$ and that $u_1(w_1, f_{\mu'(1)}, \mu'(1)) > u_1(w_1, f_{\mu_{\mathbf{w}}(1)}, \mu_{\mathbf{w}}(1))$. But then, as $\mu_{\mathbf{w}}$ is worker optimal at \mathbf{w} , firm $\mu'(1)$ is not achievable

³⁰ If a worker-optimal matching $\mu_{\mathbf{w}'}$ exists at \mathbf{w}' then $\mu' = \mu_{\mathbf{w}'}$.

for worker 1 at any complete-information stable matching at \mathbf{w} . Let \hat{i} be the worker that $\mu'(1)$ is matched with at $\mu_{\mathbf{w}}$. As $\mu'(1)$ is not achievable for worker 1 at \mathbf{w} , we have $v_{\mu'(1)}(w_{\hat{i}}, f_{\mu'(1)}, \hat{i}) \geq v_{\mu'(1)}(w_1, f_{\mu'(1)}, 1)$; otherwise, firm $\mu'(1)$ and worker 1 would block $\mu_{\mathbf{w}}$ at \mathbf{w} . Further, as there is strict preference at \mathbf{w} , we must have $v_{\mu'(1)}(w_{\hat{i}}, f_{\mu'(1)}, \hat{i}) > v_{\mu'(1)}(w_1, f_{\mu'(1)}, 1)$.

After the matching μ' is implemented, firm $\mu'(1)$ learns that worker 1's type is w_1 and not w'_1 . Moreover, firm $\mu'(1)$ strictly gains from reporting worker 1's true type w_1 because then $\mu_{\mathbf{w}}$ is implemented and $\mu'(1)$ is matched with \hat{i} . Hence, worker 1 does not profit from the deviation.

Next, suppose that after firms learn the types of the workers they are matched with at $\mu_{\mathbf{w}}$, firm j incorrectly claims that the worker $v_{\mathbf{w}}(j)$ lied about his type. But this does not change the utility that any other worker derives from matching with firm j and therefore, because $\mu_{\mathbf{w}}$ is stable, does not lead to a blocking pair with firm j and some other worker. Hence, firm j cannot benefit from this misreport. \square

References

- Bernheim, D., 1984. Rationalizable strategic behavior. *Econometrica* 52, 1007–1028.
- Chakraborty, A., Citanna, A., Ostrovsky, M., 2010. Two-sided matching with interdependent values. *J. Econ. Theory* 145, 85–105.
- Dizdar, D., Moldovanu, B., 2013. On the importance of uniform sharing rules for efficient matching. *J. Econ. Theory* 165, 106–123.
- Dubins, L., Freedman, D., 1981. Machiavelli and the Gale–Shapley Algorithm. *Am. Math. Mon.* 88, 485–494.
- Dutta, B., Vohra, R., 2005. Incomplete information, credibility, and the core. *Math. Soc. Sci.* 50, 148–165.
- Ehlers, L., Masso, J., 2007. Incomplete information and singleton cores in matching markets. *J. Econ. Theory* 36, 587–600.
- Forges, F., 2004. The ex ante incentive compatible core of the assignment game. *Math. Soc. Sci.* 47, 135–151.
- Forges, F., Minelli, E., Vohra, R., 2002. Incentives and the core of an exchange economy. *J. Math. Econ.* 38, 1–41.
- Gale, D., Shapley, L., 1962. College admissions and the stability of marriage. *Am. Math. Mon.* 69, 9–15.
- Gale, D., Sotomayor, M., 1985. Ms. Machiavelli and the stable matching problem. *Am. Math. Mon.* 92, 261–268.
- Holmstrom, B., Myerson, R., 1983. Efficient and durable decision rules with incomplete information. *Econometrica* 51, 1799–1819.
- Lazarova, E., Dimitrov, D., 2013. Paths to stability in two-sided matching under uncertainty. *Int. J. Game Theory*. <http://dx.doi.org/10.1007/s00182-015-0519-1>, in press.
- Liu, Q., Mailath, G., Postlewaite, A., Samuelson, L., 2014. Stable matching with incomplete information. *Econometrica* 82, 541–587.
- Majumdar, D., 2003. Ordinally Bayesian Incentive Compatible Stable Matchings. Working paper.
- Pearce, D., 1984. Rationalizable strategic behavior and the problem of perfection. *Econometrica* 52, 1029–1050.
- Pomatto, L., 2015. Stable Matching Under Forward-Induction Reasoning. Working paper.
- Roth, A., 1982. The economics of matching: stability and incentives. *Math. Oper. Res.* 7, 617–628.
- Roth, A., 1984. Misrepresentation and stability in the marriage problem. *J. Econ. Theory* 34, 383–387.
- Roth, A., 1989. Two-sided matching with incomplete information about others' preferences. *Games Econ. Behav.* 1, 191–209.
- Roth, A., Sotomayor, M., 1990. *Two-Sided Matching*. Cambridge University Press, Cambridge, UK.
- Vohra, R., 1999. Incomplete information, incentive compatibility, and the core. *J. Econ. Theory* 86, 123–147.
- Wilson, R., 1978. Information, efficiency, and the core of an economy. *Econometrica* 46, 807–816.
- Yenmez, B., 2013. Incentive-compatible matching mechanisms: consistency with various stability notions. *Am. Econ. J. Microecon.* 5, 120–141.