A Theory of Stability in Matching
with Incomplete Information*
(Preliminary)

Yi-Chun Chen† and Gaoji Hu ‡

December 16, 2017

Abstract

This paper provides a framework to study two-sided matching markets with incomplete information. In those markets, when two agents from opposite sides evaluate each other, they take their opponent’s evaluation into account, which initiates agents’ higher order reasoning. We propose a blocking notion that captures this higher order reasoning and then a stability notion that captures, in addition to individual rationality and no blocking, the idea that absence of blocking conveys no further information. We show that starting from an arbitrary allocation and an arbitrary information structure, the process of allowing randomly chosen blocking pairs to rematch, accompanied with information updating, will converge to an allocation that is stable under the updated information structure with probability one. We study the welfare effect when there is entry to a stable market. We show that an added agent will make the welfare of all other agents on the same side lower and that of all agents on the opposite side higher in new market states that are stable, when the original market state is complete-information stable. Counterexamples are provided when this condition is not satisfied. Our framework also facilitates (i) the study of alternative stability notions such as ex ante stability and Bayesian stability and (ii) the study of models with richer details such as partially verifiable types.

JEL Classification: C62, C78, D83

Keywords: two-sided matching, incomplete information, stability, convergence, comparative statics, Bayesian stability.

*We are grateful to Fuhito Kojima for his encouragement and many helpful discussions. We also thank Bo Chen, Laura Doval, Jiangtao Li, Qingmin Liu, Xiao Luo, George Mailath, Mike Ostrovsky, Andy Postlewaite, Ning Sun, Xiang Sun, Satoru Takahashi and the audience at the International Conference on Game Theory & the Fourth Microeconomics Workshop (Nanjing) for their helpful comments and suggestions. Part of this paper was written while Hu was visiting Stanford University and he would like to thank the institution for its hospitality and support. This work is supported by Singapore MOE AcRF Tier 1 Grant. All remaining errors are our own.

†Department of Economics, National University of Singapore, S(117570). ecsycc@nus.edu.sg.
‡Department of Economics, National University of Singapore, S(117570). hu.gaoji@gmail.com.
1 INTRODUCTION

Stability of matchings in two-sided markets has been connected to both equity and efficiency, two of most important objectives in economics. A prevailing assumption in this literature is that the information is complete, i.e., the characteristics and preferences of all market participants are publicly known. However, incomplete information is ubiquitous in matching markets: a man may have imprecise information about his girlfriend, which might be learnt after they get married; firms may not know the productivity of their potential employees before employment; professors may recruit incompetent PhD students based on their undergraduate transcripts; and etc.

We consider a matching model where workers are matched with firms and depart from the prevailing assumption of complete information. We first describe how agents form and update their information about all market participants’ types, and then propose an incomplete-information stability notion that allows for flexible information structure. Our formulation facilitates the study of matching processes, particularly the path to stability problem. Conceptually, our formulation also facilitates the study of other stability notions in matching with incomplete information, such as ex ante stability and Bayesian stability.

Stability under complete information requires individual rationality, such that each agent has a nonnegative payoff, and no blocking pair; in the present context, no blocking pair means that no worker-firm pair would prefer being matched with each other at a certain wage to staying with the current matching. In contrast, in a job market with incomplete information, a typical agent may not know the types of their potential employees/employers which reflect their productivity/desirability. Without the information, however, the firm and the worker do not know if they would prefer each other to their current partners. As a result, the notion of blocking and stability in the complete-information environment is no longer appropriate.

In the incomplete-information setting, each agent is associated with a type, which determines agents’ payoffs from a match. Moreover, an agent’s information is described by a partition over possible type profiles. We assume that agents’ partitions are common knowledge, that agents who are uncertain about their potential partners’ types care about

---

1See Balinski and Sönmez (1999) and Abdulkadiroğlu and Sönmez (2003) for how stability implies the elimination of justified envy, a basic fairness axiom. See Shapley and Shubik (1971) and Liu et al. (2014) for how stability leads to efficiency.

2Knuth (1976) provides an example of a blocking path that admits a cycle, i.e. any matching on the path is not stable. This motivates the study of the convergence of blocking paths. Pursuing an affirmative answer to the above question results in a fruitful literature. See Roth and Vande Vate (1990), Kojima and Ünver (2008), Klaus and Klijn (2007), Chen et al. (2016), Fujishige and Yang (2016) and Chen and Hu (2017).
the worst possibility, and that an agent can observe the type of her/his own partner.\(^3\) A market state in our setting specifies an allocation (i.e., a matching together with a prevailing wage profile), a realized type profile, and a profile of partitions which represents agents’ flexible and heterogeneous information about the realized type profile. Stability of a state captures the idea that the allocation is individually rational and admits no blocking pair with respect to the information structure; and moreover, the absence of blocking conveys no further information to the agents. The last requirement, in particular, formalizes “informational stability,” which is specific to the incomplete information setting.

Equipped with the notion of stability, we then study a matching process which mimics the individual behavior of searching for desired job, school and life partner. Indeed, if a worker and a firm find themselves better off being matched with each other than in maintaining the status quo, they will form a match together to make an improvement. The new matching may again admit a blocking pair and thus a rematching opportunity, which results in another new matching, and so on. One prominent question is whether such a process finally stops at a stable matching.

When information is incomplete, along the matching process each observation of rematching or no rematching provides information to both firms and workers. Consequently, even if no rematching is observed under an information structure, there may be some rematching opportunity when information is updated. Therefore, a matching process is necessarily associated with an information updating process where agents draw inference from each observation. For an arbitrary initial market state, this learning and rematching process consists of a sequence of states. We call it a \textit{Learning-Blocking Path}. We show that when blocking pairs are randomly selected to rematch, the resulting \textit{Learning-Blocking Path} converges to a stable state with probability one.

Given a stable market state, new workers or new positions may initiates \textit{Learning-Blocking Paths}. We study the welfare effect of adding one worker or one firm into a stable market state. In the complete information setting with non-transferable utility, a new worker entering into a stable market makes existing workers worse off and existing firms better off; symmetrically, a new firm entering into a stable market state makes existing firms worse off and existing workers better off.\(^4\) With incomplete information, however, these properties may fail due to correlation between agents’ types.\(^5\) Nevertheless, if we consider a restricted domain, i.e., when the original market allocation is not only

\(^3\)See Liu et al. (2014), Bikhchandani (2017) and Chen and Hu (2017) for similar assumptions. We revisit the last two assumptions in Section 8 and Section 9, respectively.

\(^4\)See, for example, Blum et al. (1997).

\(^5\)Moreover, the set of stable matchings does not necessarily form a lattice and the Rural Hospital Theorem fails. See, for example, Roth and Sotomayor (1990) for results of the lattice structure (pp. 60 and pp. 208) and the Rural Hospital Theorem (pp. 144).
stable but also complete-information stable, then the comparison between agents’ welfare under the (incomplete-information) stable states in the new market and their welfare under the stable state in the original market can be restored.

Our formulation also contributes to the literature conceptually, in the sense that it facilitates the study of other stability notions. The first notion we investigate is ex ante stability. It describes what kind of matching outcomes (i.e., an allocation and a realized type profile) could arise when agent-specific information structure is not realized /specified. Initially, the candidate outcomes include all individually rational ones. Then the outcomes that are surely blocked will be iteratively eliminated. The elimination procedure converges to a set of outcomes which we refer as the set of ex ante stable matching outcomes. We show the intuitive result that when we extract the allocations from all stable states, it is exactly the set of allocations we can extract from the set of ex ante stable outcomes. We finally study the notion of Bayesian stability where within a potential blocking pair, the firm and the worker evaluate each other based on the expected payoffs. Our analysis is not limited to a specific environment. The idea can be applied to a broader class of models. For example, we can consider an agent’s type as of a richer structure. Each type consists of many components. Each component is either unobservable to others in any case, or observable to a specific agent when matched.

The rest of this section reviews the related literature. Section 2 introduces the model. Section 3 defines blocking and stability with incomplete information. Section 4 revisits our blocking criteria from different angles. The path to stability problem is studied in Section 5. Adding agents into a stable market state may trigger Learning-Blocking Paths. Section 6 investigates the welfare effect of adding agents. Sections 7 and 8 study two alternative stability notions, namely ex ante stability and Bayesian stability. Section 9 concludes, where we discuss the generality of our analysis.

**Related Literature**

The seminal model of Gale and Shapley (1962), studying the marriage market (college admission market), has been used in a large literature of two-sided matching. Many classical theories are surveyed in Roth and Sotomayor (1990) and more recently by, for example, Roth (2008), Abdulkadiroglu and Sönmez (2013) and Kojima et al. (2017). In this literature, a prevailing assumption is that the information is complete.

Recently, Liu et al. (2014) (hereafter, LMPS) studied stable matchings with one-sided incomplete information. Particularly, in the job market setting, workers have private types while firms possess some information about worker-types. They introduces an ex
uate notion of stability to predict what kind of matching outcomes could arise when heterogeneous firm-specific information is not realized(specified). In a previous paper, Chen and Hu (2017) (hereafter, CH), we formulate heterogeneous firm-specific information and studied a stability notion that is defined with respect to an information structure. In both papers, firms who are uncertain about a worker’s type use the worst-case desideratum to evaluate him. The current paper extends the stability notions of both LMPS and CH to the setting with two-sided incomplete information, and relaxes the worst-case desideratum to study Bayesian stability.

The key subtlety when both sides have incomplete information, compared with the one-sided incomplete-information setting discussed in LMPS and CH, is that the worst-case desideratum of a potential blocking pair involves interaction between the willingness of the worker and the firm. More precisely, when the potential blocking firm thinks about the worst possible type of the potential blocking worker, she takes into account the blocking worker’s perception of the worst possible type of the blocking firm herself, which in turn takes into account herself’s perception of the worst possible type of the blocking worker, etc. The potential blocking pair run through such an iterative thinking until any higher level of reasoning provides no help to their evaluation of each other.

Whether or not a matching process converges to a stable allocation is known as the paths to stability problem. See CH for the detailed discussion of convergent LEARNING-BLOCKING PATHs in matching with one-sided incomplete information, and references there for related studies. The welfare effect of adding agents is a standard comparative statics analysis. See Theorem 2.26 in Roth and Sotomayor (1990) for a textbook treatment, and Blum et al. (1997) for further studies.

The stability notion of Bikhchandani (2017) is similar to that of LMPS but focuses on non-transferable utility and Bayesian stability, instead of the worst-case desideratum. Pomatto (2015) considers a matching game and derives the same set of incomplete-information stable outcomes as LMPS by using forward induction reasoning. In a model with non-transferable utility, Lazarova and Dimitrov (2013) study stability under a permissive notion of blocking, which is not applicable when the blocking notions of LMPS, CH, Bikhchandani (2017), Pomatto (2015), and the current paper are adopted.

Chakraborty et al. (2010) studies college admission markets where the students’ characteristics are unobservable for colleges but colleges may receive noisy signals about these characteristics. They proposed a stable matching mechanism which specifies a matching and how much information to reveal to the agents, given each reported signal profile with an information structure. They also made clear that stability of a matching
mechanism should depend on the agents’ information structure.\textsuperscript{6} In a similar vein, we define a stable allocation with respect to different information structures of the agents.

Another stream of literature studies agents’ incomplete information about other’s preferences. Starting from Roth (1989), this literature includes the results about stable mechanisms and manipulation of preferences (see, for example, Ehlers and Massô (2007, 2015) and Yenmez (2013)). Our paper differs from them in at least two aspects. First, an agent’s type in our model represents both her/his incomplete information about other’s characteristics (e.g. workers having the same preference over firms may or may not have the same productivity, where productivity is interpreted as characteristics) and about others’ preferences, instead of only preferences.\textsuperscript{7} Second, they focus on stability of (in)direct matching mechanisms while we focus on stable matchings.

The literature of core is related to this paper for similar reasons as LMPS. In our context, a coalition is simply a worker-firm pair. Wilson (1978) proposed two polar cases about what information a coalition might use in evaluating outcomes and formulating objections: sharing all information or only the common knowledge information. Different information-sharing assumptions at the interim stage may determine different blockings to a potential core allocation. Dutta and Vohra (2005) considered a scenario between the two polar cases where agents are able to make conditional offers to other agents that have no effect unless the offer is accepted by the other agents. The interim no blocking notion introduced by Yenmez (2013) is similar to this but in matching environment. Although our blocking notion shares similar spirit, our stability notion capture more important features that are different from Dutta and Vohra (2005). Namely, Dutta and Vohra (2005) predicts an allocation in core if it is immune to blocking; while we, as LMPS, predict an allocation in core if it is not only immune to blocking, but also immune to blockings caused by the information that can be inferred from being immune to blocking and similar higher order reasoning.

2 The Model

We consider the following setup of matching with incomplete information. The setup generalizes the complete-information matching models studied by Shapley and Shubik (1971) and Crawford and Knoer (1981), and further the incomplete-information models by Liu et al. (2014), Pomatto (2015), Bikhchandani (2017) and Chen and Hu (2017).

\textsuperscript{6}To the best of our knowledge, Chakraborty et al. (2010) is the first paper that internalizes available information into stability notions.

\textsuperscript{7}One agent’s preference is over other agents’ characteristics. Thus uncertainty about others’ characteristics can also be interpreted as uncertainty about agents’ own preferences.
There is a finite set $I$ of workers to be matched with a finite set $J$ of firms. Denote a generic worker by $i$ and a generic firm by $j$. Each agent’s index is publicly observed. The productivity of an agent is however determined by the agent’s type. Let $W$ be the finite set of possible worker types and $F$ be the finite set of possible firm types. The type assignment function is denoted by $\mathbf{t} := (\mathbf{w}, \mathbf{f})$, where $\mathbf{w} : I \rightarrow W$ assigns each worker a type and $\mathbf{f} : J \rightarrow F$ assigns each firm a type. We denote by $T$ a set of type assignment functions, i.e., $T \subset W^{|I|} \times F^{|J|}$.

A match between worker type $w \in W$ and firm type $f \in F$ gives rise to the 
worker premuneration value $\nu_{wf} \in \mathbb{R}$ and firm premuneration value $\phi_{wf} \in \mathbb{R}$. The sum of the premuneration values $\nu_{wf} + \phi_{wf}$ is called the surplus of the match. Denote these values by $\nu_{w(\emptyset),f(j)}$ for the unmatched worker and $\phi_{w(i),f(\emptyset)}$ for the unmatched firm, both of which are set to be zero.

Given a match between worker $i$ (of type $\mathbf{w}(i)$) and firm $j$ (of type $\mathbf{f}(j)$), the worker’s payoff and the firm’s payoff are, respectively, $\nu_{\mathbf{w}(i),\mathbf{f}(j)} + p$ and $\phi_{\mathbf{w}(i),\mathbf{f}(j)} - p$, where $p \in \mathbb{R}$ is the payment made to worker $i$ by firm $j$.

A matching is a function $\mu : I \rightarrow J \cup \{\emptyset\}$, one-to-one on $\mu^{-1}(J)$, that assigns worker $i$ to firm $\mu(i)$, where $\mu(i) = \emptyset$ means that worker $i$ is unemployed and $\mu^{-1}(j) = \emptyset$ means that firm $j$ does not hire a worker.

A payment scheme $\mathbf{p}$ associated with a matching $\mu$ is a vector that specifies a payment $p_{i,\mu(i)} \in \mathbb{R}$ for each $i \in I$ and $p_{\mu^{-1}(j),j} \in \mathbb{R}$ for each $j \in J$. To avoid nuisance cases, we associate zero payments with unmatched agents, setting $p_{\emptyset,j} = p_{i,\emptyset} = 0$.

An allocation $(\mu, \mathbf{p})$ consists of a matching $\mu$ and an associated payment scheme $\mathbf{p}$. We assume that the entire allocation is publicly observable. Denote by $\mathcal{A}$ the set of allocations.

Each agent knows her/his own type, but other agents only know that the type assignment function $\mathbf{t}$ belongs to a set $T$. The agents may have imprecise or wrong information about the realized type assignment function. A generic agent $k$’s information is described by a partition $\Pi_k$ of $T$, where $k \in I \cup J$. For any $\mathbf{t} \in T$, we write $\Pi_k(\mathbf{t})$ as the element of the partition that contains $\mathbf{t}$. Each $\mathbf{t}' \in \Pi_k(\mathbf{t}) \in \Pi_k$ is a possible type assignment from agent $k$’s viewpoint when the realized type assignment function is $\mathbf{t}$. Denote the profile of partitions by $\Pi$, i.e., $\Pi := (\{\Pi_i\}_{i \in I}, \{\Pi_j\}_{j \in J})$, which is common knowledge among agents. Following LMPS, we assume that agents can observe the type of their partners, i.e., within a matched pair, the firm and the worker can observe the type of each other. Finally, the functions $\nu : W \times F \rightarrow \mathbb{R}$ and $\phi : W \times F \rightarrow \mathbb{R}$ are common knowledge.

---

8See Mailath et al. (2012, 2013) for discussions on premuneration values.
A state of the matching market, \((\mu, p, t, \Pi)\), specifies an allocation \((\mu, p)\), a type assignment function \(t\) and a partition profile \(\Pi\) such that each agent’s partition is consistent with her/his own type and her/his partner’s type observed by her/him, i.e., for every \(i \in I\) and every \(t' \in T\), \(t''(i) = t'(i)\) and \(t''(\mu(i)) = t'(\mu(i))\) for all \(t'' \in \Pi_i(t')\); for every \(j \in J\) and every \(t' \in T\), \(t''(j) = t'(j)\) and \(t''(\mu^{-1}(j)) = t'(\mu^{-1}(j))\) for all \(t'' \in \Pi_j(t')\).

### 3 Stability with Incomplete Information

#### 3.1 Individual Rationality

A state is said to be individually rational if each agent receives at least the payoff given by the outside option of remaining unmatched.

**Definition 1.** A state \((\mu, p, t, \Pi)\) is said to be individually rational if

\[
\nu_{t(i), t(\mu(i))} + p_{i, \mu(i)} \geq 0 \text{ for all } i \in I \text{ and } \\
\phi_{t(\mu^{-1}(j)), t(j)} - p_{\mu^{-1}(j), j} \geq 0 \text{ for all } j \in J.
\]

#### 3.2 Blocking

The notion of incomplete-information “blocking” naturally extends its complete-information counterpart. In particular, a matching is blocked if some worker-firm pair \((i, j)\) can mutually benefit from matching with each other. Following LMPS, we assume that a firm is concerned about the worst case of a worker, as she does not know his true type; similarly, a worker is concerned about the worst case of a firm.\(^9\) We assume that this worst-case desideratum is common knowledge.

Consider a potential “blocking pair” \((i, j)\) for the state \((\mu, p, t, \Pi)\). A type assignment function is relevant for worker \(i\) only when firm \(j\) is willing to participate the blocking. Thus it is sufficient for worker \(i\) to participate the blocking if he can guarantee an improvement for those type assignments in \(\Pi_i(t)\) such that firm \(j\) may obtain higher payoff. Formally, when worker \(i\) evaluates a potential blocking firm \(j\) at some potential salary \(p\), a type assignment function \(t'\) is relevant for worker \(i\) if

\[
t' \in \Pi_i(t) \text{ and } \max_{t'' \in \Pi_j(t')} \left[ \phi_{t''(i), t''(j)} - p_{\mu^{-1}(j), j} \right] > 0,
\]

i.e., firm \(j\) may better off under \(t'\). All type assignment functions violating the inequality

\(^9\)See Section 8 for the discussion on Bayesian stability, where agents care about the expected payoff instead of worst-case payoff.
are irrelevant because the “blocking pair” can never be formed due to the firm’s objection: she cannot benefit even in the best case. Similarly, when firm \( j \) evaluates worker \( i \) at the salary \( p \), a type assignment \( t' \) is relevant for firm \( j \) if

\[
t' \in \Pi_j(t) \text{ and } \max_{v' \in \Pi_j(t')} [\nu_{v''(i),v''(j)} + p] - [\nu_{v''(i),v''(\mu(i))} + p_{i,\mu(i)}] > 0.
\]

(2)

Since agents commonly know that others care only about the relevant type assignment functions, potential blocking agents need to guarantee an improvement only for even less type assignment functions than specified in (1)-(2). Generally, sophisticated agents will refine the set of relevant type assignments whenever possible. Define \( \Pi^0 = \Pi \) and recursively for \( l = 1, 2, \ldots \) that

\[
\Pi_i^{[l]}(t') := \left\{ t'' \in \Pi_i(t') : \Pi_i^{[l-1]}(t'') \neq \emptyset \text{ and } \max_{\tilde{t} \in \Pi_i^{[l-1]}(t'')} [\phi_{\tilde{t}(i),\tilde{t}(j)} - p] - [\phi_{\tilde{t}(\mu^{-1}(j)),\tilde{t}(j)} - p_{\mu^{-1}(j),j}] > 0 \right\}
\]

(3)

\[
\Pi_j^{[l]}(t') := \left\{ t'' \in \Pi_j(t') : \Pi_i^{[l-1]}(t'') \neq \emptyset \text{ and } \max_{\tilde{t} \in \Pi_i^{[l-1]}(t'')} [\nu_{\tilde{t}(i),\tilde{t}(j)} + p] - [\nu_{\tilde{t}(\mu),\tilde{t}(\mu)} + p_{i,\mu}] > 0 \right\},
\]

(4)

where \( t' \in T \) and, for simplicity, the dependence of \( \Pi_i^{[l]} \) and \( \Pi_j^{[l]} \) on \((\mu, p, \Pi)\) and \((i, j; p)\) is suppressed in the notation.

Since \( \Pi_j^{[1]}(t'') \subset \Pi_j^{[0]}(t'') \) for all \( t'' \in T \), we know that \( \Pi_j^{[1]}(t'') \neq \emptyset \) only if \( \Pi_j^{[0]}(t'') \neq \emptyset \) and that

\[
\max_{\tilde{t} \in \Pi_j^{[1]}(t'')} [\phi_{\tilde{t}(i),\tilde{t}(j)} - p] - [\phi_{\tilde{t}(\mu^{-1}(j)),\tilde{t}(j)} - p_{\mu^{-1}(j),j}] > 0
\]

only if

\[
\max_{\tilde{t} \in \Pi_j^{[0]}(t'')} [\phi_{\tilde{t}(i),\tilde{t}(j)} - p] - [\phi_{\tilde{t}(\mu^{-1}(j)),\tilde{t}(j)} - p_{\mu^{-1}(j),j}] > 0.
\]

Therefore, \( \Pi_i^{[2]}(t') \subset \Pi_i^{[1]}(t') \) for all \( t' \in T \). Similarly, \( \Pi_j^{[2]}(t') \subset \Pi_j^{[1]}(t') \) for all \( t' \in T \). By induction, \( \Pi_k^{[l]}(t') \) is decreasing in \( l \) for all \( t' \in T \) and \( k \in I \cup J \). Let \( l^* \) be a sufficiently large integer such that \( \Pi_k^{[l^*+1]}(t') = \Pi_k^{[l^*]}(t') \) for all \( t' \in T \) and \( k = i, j \).

If \( \Pi_i^{[l]}(t') = \emptyset \) for some \( l \) and some \( t' \in T \), then worker \( i \) thinks that a rational firm \( j \) would never prefer worker \( i \) at the salary \( p \), were \( t' \) the true type assignment function. Furthermore, if \( \Pi_i^{[l^*]}(t) = \emptyset \), then worker \( i \) thinks that a rational firm \( j \) would never prefer worker \( i \) at the salary \( p \). In these cases, worker \( i \) would never prefer firm \( j \) at the salary \( p \) as well. Symmetric argument applies for firms.
Definition 2. A state \((\mu, p, t, \Pi)\) is said to be blocked if there exists a worker-firm pair \((i, j)\) and a payment \(p \in \mathbb{R}\) such that \(\Pi_i^{[r]}(t) \neq \emptyset\), \(\Pi_j^{[r]}(t) \neq \emptyset\) and

\[
\nu_{\psi(i), t(j)} + p > \nu_{\psi(i), t(\mu(i))} + p_{t, \mu(i)} \text{ for all } t' \in \Pi_i^{[r]}(t) \text{ and } \phi_{t(i), t(j)} - p > \phi_{t(i), t(j)} - p_{t^{-1}(j), j} \text{ for all } t' \in \Pi_j^{[r]}(t). \tag{5} \tag{6}
\]

The tuple \((i, j; p)\) is called a blocking combination. The following example illustrates the iteration of (3)-(4), which is nontrivial for natural environments, i.e., one-dimensional type and premuneration functions of product form.

Example 1. Consider a market with two workers, \(I = \{x, y\}\), and two firms, \(J = \{a, b\}\). The set of possible type assignment functions is given by \(T = \{t^1, t^2, t^3, t^4\}\), where \(t^n, n = 1, \ldots, 4\), are listed below.

<table>
<thead>
<tr>
<th></th>
<th>x</th>
<th>y</th>
<th>a</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t^1)</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>(t^2)</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>(t^3)</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>(t^4)</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

The premuneration values for workers and firms are given by the product form, i.e., \(\nu_{w,f} = \phi_{w,f} = w.f\). Suppose \(t^1\) is the realized type assignment function. Consider a status quo matching \(\mu\) with \(\mu(x) = a\) and \(\mu(y) = b\), where the payments are all zero, and consider a potential blocking combination by \((a, y; 0)\). We list the correspondences \(\Pi_i^{[l]}\) and \(\Pi_j^{[l]}\) in Table 1 for \(l = 0, 1, \ldots, 3\). Note that \(\Pi_i^{[0]}\) and \(\Pi_j^{[0]}\) are partial information functions while others may not be. In addition, \(\Pi_i^{[4]} = \Pi_j^{[3]}\) and \(\Pi_j^{[4]} = \Pi_j^{[3]}\). Obviously, both \(\Pi_i^{[3]}(t^1)\) and \(\Pi_j^{[3]}(t^1)\) are nonempty. Since conditions (5)-(6) are satisfied, we know that the status quo state is blocked by \((a, y; 0)\).

Recall that complete-information blocking notion conveys the intuition that the blocking pair could improve their payoffs by rematching with each other. Since \(\Pi_i^{[r]}(t)\) and \(\Pi_j^{[r]}(t)\) are only subsets of \(\Pi_i(t)\) and \(\Pi_j(t)\) respectively, they may not include the realized type profile \(t\). Therefore, even if (5)-(6) are satisfied, rematching may not result in higher payoffs for the blocking pair, where payoffs are evaluated at the realized types. To see that (5)-(6) actually guaranteed an improvement for both agents under the realized type profile, we prove the following lemma.

Lemma 1. If \((\mu, p, t, \Pi)\) is blocked, then

\[
\nu_{t(i), t(j)} + p > \nu_{t(i), t(\mu(i))} + p_{t, \mu(i)} \text{ and } \phi_{t(i), t(j)} - p > \phi_{t(\mu^{-1}(j)), t(j)} - p_{t^{-1}(j), j}.\]
Table 1: The correspondences $\Pi_a^l$ and $\Pi_y^l$ for $l = 0, 1, 2, 3$. 

<table>
<thead>
<tr>
<th></th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$t_3$</th>
<th>$t_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Pi_a^0$</td>
<td>${t_1, t_2}$</td>
<td>${t_1, t_2}$</td>
<td>${t_2, t_4}$</td>
<td>${t_3, t_4}$</td>
</tr>
<tr>
<td>$\Pi_y^0$</td>
<td>${t_1}$</td>
<td>${t_2, t_3}$</td>
<td>${t_2, t_3}$</td>
<td>${t_4}$</td>
</tr>
<tr>
<td>$\Pi_a^1$</td>
<td>${t_1, t_2}$</td>
<td>${t_1, t_2}$</td>
<td>${t_3}$</td>
<td>${t_3}$</td>
</tr>
<tr>
<td>$\Pi_y^1$</td>
<td>${t_1}$</td>
<td>${t_2, t_3}$</td>
<td>${t_2, t_3}$</td>
<td>${t_4}$</td>
</tr>
<tr>
<td>$\Pi_a^2$</td>
<td>${t_1, t_2}$</td>
<td>${t_1, t_2}$</td>
<td>${t_3}$</td>
<td>${t_3}$</td>
</tr>
<tr>
<td>$\Pi_y^2$</td>
<td>${t_1}$</td>
<td>${t_2}$</td>
<td>${t_2}$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>$\Pi_a^3$</td>
<td>${t_1}$</td>
<td>${t_1}$</td>
<td>$\emptyset$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>$\Pi_y^3$</td>
<td>${t_1}$</td>
<td>${t_2}$</td>
<td>${t_2}$</td>
<td>$\emptyset$</td>
</tr>
</tbody>
</table>

We now provide a useful lemma, showing the intuition that more precise information makes it easier to form a blocking. Formally, we say a partition $\Pi_k$ for an agent $k$ is finer than $\Pi'_k$ (equivalently, $\Pi'_k$ is coarser than $\Pi_k$) if for each $t' \in T$, $\Pi_k(t') \subset \Pi'_k(t')$. We say a partition profile $\Pi$ is finer than $\Pi'$ (equivalently, $\Pi'$ is coarser than $\Pi$) if $\Pi_k$ is finer than $\Pi'_k$ for all $k \in I \cup J$.

**Lemma 2.** Suppose $(\mu, p, t, \Pi')$ and $(\mu, p, t, \Pi)$ are two states such that $\Pi'$ is coarser than $\Pi$. If $(\mu, p, t, \Pi')$ is blocked, then $(\mu, p, t, \Pi)$ is blocked.

### 3.3 Stability

When $\Pi_k(t') = \{t'\}$ for all $k \in I \cup J$ and all $t' \in T$, i.e., information is complete, a stable state is simply a state that is individually rational and not blocked. The notion of incomplete-information “stability” differs from its complete-information counterpart in the following way: When information is complete, stable matching conveys the intuition that when “the agents have a very good idea of one another’s preferences and have easy access to each other, ... we might expect that stable matchings will be especially likely to occur” (Roth and Sotomayor, 1990, pp. 22). In contrast, the partition $\Pi_k$ in our setting describes only agent $k$’s imprecise idea about the other agents’ information. The “stability” notion that we propose below captures: i) the state is individually rational, ii) the state is not blocked, and, most importantly, iii) it is common knowledge that the state is individually rational and not blocked. The first two properties coincide with the complete-information stability, while the third one describes “information stability” which is specific to the incomplete-information environment.
Given a state \((\mu, p, t, \Pi)\), let \(N(\mu, p, \Pi)\) be a binary partition of \(T\), depending on the public information \((\mu, p, \Pi)\) of a state, such that \(N(\mu, p, \Pi)(t') = N(\mu, p, \Pi)(t'')\) if and only if either neither \((\mu, p, t', \Pi)\) nor \((\mu, p, t'', \Pi)\) is blocked or both of them are blocked (by a pair or an individual). To formally describe “information stability”, we define the following allocation dependent operator \(H_{\mu, p}\) that maps one partition profile \(\Pi\) to another (the meet of \(\Pi_k\) and \(N(\mu, p, \Pi)\) for each agent \(k \in I \cup J\)):

\[
[H_{\mu, p}(\Pi)]_k(t') := \Pi_k(t') \cap N(\mu, p, \Pi)(t') \quad \text{for all} \quad t' \in T \quad \text{and all} \quad k \in I \cup J.
\]

To simplify the notation, we denote the meet of two partitions (a partition profile and another partition) by \(\wedge\), i.e., \(H_{\mu, p}(\Pi) = \Pi \wedge N(\mu, p, \Pi)\). A state is said to be stable if it is individually rational, not blocked, and no information can be inferred from the fact that the state is individually rational and not blocked.

**Definition 3.** A state \((\mu, p, t, \Pi)\) is said to be stable if

1. it is individually rational,
2. it is not blocked, and
3. \(\Pi\) is a fixed point of \(H_{\mu, p}\), i.e. \(H_{\mu, p}(\Pi) = \Pi\).

As in LMPS and CH, the existence of a stable state is guaranteed by the existence of a complete-information stable state, i.e., if \(\Pi_k(t') = \{t'\}\) for all \(k \in I \cup J\) and \(t' \in T\), then for each \(t\) there exists an allocation \((\mu, p)\) such that the state \((\mu, p, t, \Pi)\) is stable (see Shapley and Shubik (1971) and Crawford and Knoer (1981)).

Given a stable state \((\mu, p, t, \Pi)\), if we blur the information, i.e., replace \(\Pi\) with a coarser partition profile, the resulted state may not be stable any more. See Example 2 below. However, the resulted state is “essentially” stable. Formally, we say that a state \((\mu, p, t, \Pi)\) is quasi-stable if \((\mu, p, t, H_{\mu, p}^k(\Pi))\) is stable for some \(k \in \mathbb{N}\).

**Example 2.** Consider a market with two workers, \(I = \{x, y\}\), and two firms, \(J = \{a, b\}\). The set of possible type assignment functions is given by \(T = \{t^1, t^2, t^3\}\), where \(t^1, t^2,\) and \(t^3\), are listed below.

\[
\begin{array}{cccc}
  x & y & a & b \\
  t^1 & 2 & 3 & 2 & 3 \\
  t^2 & 2 & 3 & 2 & 1 \\
  t^3 & 4 & 3 & 2 & 3 \\
\end{array}
\]

The premuneration values for workers and firms are both given by product of types, i.e., \(\nu_{w,f} = \phi_{w,f} = wf\). The realized type assignment function is \(t^1\). Consider a matching \(\mu\) with \(\mu(x) = a\) and \(\mu(y) = b\). The payments are all zero, i.e., \(p = 0\).
For the complete-information partition profile $\Pi$, i.e., $\Pi_k = \{\{t^1\}, \{t^2\}, \{t^3\}\}$ for all $k \in I \cup J$, the state $(\mu, p, t^1, \Pi)$ is stable. However, when we consider an alternative $\Pi'$ below, $(\mu, p, t^1, \Pi')$ is not stable because $H_{\mu, p}(\Pi') = \Pi \neq \Pi'$, where $\Pi'$ is listed below. Nevertheless, $(\mu, p, t^1, H_{\mu, p}(\Pi'))$ is stable.

$$
\begin{align*}
\Pi_x' &= \{\{t^1, t^2\}, \{t^3\}\} \\
\Pi_y' &= \{\{t^1\}, \{t^2\}, \{t^3\}\}
\end{align*}
$$

**Proposition 1.** Suppose $(\mu, p, t, \Pi')$ and $(\mu, p, t, \Pi)$ are two states such that $\Pi'$ is coarser than $\Pi$. If $(\mu, p, t, \Pi)$ is stable, then $(\mu, p, t, \Pi')$ is quasi-stable.\(^{10}\)

**Remark 1.** Instead of a fixed point property, we can also interpret a stable state as a result of iteration. For example, suppose $(\mu, p)$ is a complete-information stable allocation when the type profile is $t$. Given an arbitrary partition profile $\Pi$, we consider the iteration with $\Pi_0 := \Pi$ and for $k = 1, 2, \ldots$, $\Pi_k = H_{\mu, p}(\Pi_{k-1})$. Then $\Pi^\infty$ is a fixed point of $H_{\mu, p}$, i.e., $(\mu, p, t, \Pi^\infty)$ is a stable state.

### 3.4 Stable Allocations

Given a type assignment function $t$, the payoffs of agents are actually determined by the allocation. Therefore, we concentrate on the set of allocations that can arise in stable states, which is a set $S$ of allocations that are stable under some partition profile, i.e.

$$
S(t) := \{ (\mu, p) \in A : \exists \Pi \text{ s.t. } (\mu, p, t, \Pi) \text{ is stable} \}.
$$

The ex ante set of stable allocations is then defined as $S := \bigcup_{t \in T} S(t)$. In Section 7, we will introduce ex ante stability. The two stability notions are consistent with each other in the allocation sense: the ex ante set of stable allocations are the same.

\(^{10}\)Proposition 1 captures the intuition that if a matching is stable under some information structure then it is also stable for every coarser information structure. In a setting different from our paper, (Chakraborty et al., 2010, Theorem 1) shows that if a matching mechanism is stable under some information structure it is also stable for every coarser information structure.
4 More on the Blocking Criteria

This section further discusses the blocking criteria in Subsection 3.2. Readers who are keen on exploring applications on matching processes and extensions of the stability notion can skip this section and go straight to Section 5.

4.1 Why not Taking the Minimum?

Consider a potential blocking combination \((i, j; p)\) for the state \((\mu, p, t, \Pi)\). Since it is publicly known that agents care about the worst case of their payoffs, agents could further restrict their attention, compared with (5)-(6). Formally, define \(\Pi^{(0)} = \Pi\) and recursively for \(l = 1, 2, \ldots\) that

\[
\Pi_i^{(l)}(t') := \left\{ t'' \in \Pi_i^{(l-1)}(t') : \Pi_j^{(l-1)}(t'') \neq \emptyset, \min_{t \in \Pi_j^{(l-1)}(t'')} [\phi_{t(i)}, t(j) - p] - [\phi_{t(\mu^{-1}(j)), t(j)} - p_{\mu^{-1}(j), j}] > 0 \right\}
\]

(9)

\[
\Pi_j^{(l)}(t') := \left\{ t'' \in \Pi_j^{(l-1)}(t') : \Pi_i^{(l-1)}(t'') \neq \emptyset, \min_{t \in \Pi_i^{(l-1)}(t'')} [\nu_{t(i)}, t(i) - p] - [\nu_{t(\mu(i))}, t(i) + p_{i, \mu(i)}] > 0 \right\},
\]

(10)

where the dependence of \(\Pi_i^{(l)}\) and \(\Pi_j^{(l)}\) on \((\mu, p, \Pi)\) and \((i, j; p)\) is suppressed in the notation. However, it is obvious that further reasoning provides no information, i.e. \(\Pi_i^{(l)} = \Pi_i^{(1)}\) for all \(l \in \mathbb{N}\) and \(l \geq 2\).

Suppose potential blocking pairs refine their consideration sets by (9)-(10). We follow the convention that a matching is blocked if some worker-firm pair \((i, j)\) can mutually benefit from matching with each other to introduce the following notion of blocking.

**Definition 4.** A state \((\mu, p, t, \Pi)\) is said to be blocked* if there exists a worker-firm pair \((i, j)\) and a payment \(p \in \mathbb{R}\) such that

**Condition (11) is necessary because of the following example.**
EXAMPLE 3. Consider a market with two workers, \( I = \{x, y\} \), and two firms, \( J = \{a, b\} \). The set of possible type assignment functions is given by \( T = \{t^1, t^2, t^3\} \), where \( t^1, t^2 \) and \( t^3 \) are listed below.

\[
\begin{array}{cccc}
x & y & a & b \\
t^1: & 2 & 4 & 2 & 4 \\
t^2: & 2 & 4 & 2 & 1 \\
t^3: & 2 & 1 & 2 & 1 \\
\end{array}
\]

The premuneration values for workers and firms are both given by product of types, i.e. \( \nu(w, f) = \phi(w, f) = w f \). The realized type assignment function is \( t^2 \). Consider a matching \( \mu \) with \( \mu(x) = b \) and \( \mu(y) = a \). The payments are all zero. If (11) is not imposed, then the state is blocked* by \( (y, b; 0) \) but worker \( y \) will get worse off if the blocking combination is satisfied.

The following example shows the difference of blocking and blocking*.

EXAMPLE 4. Consider a market with two workers, \( I = \{x, y\} \), and two firms, \( J = \{a, b\} \). The set of possible type assignment functions is given by \( T = \{t^1, t^2, t^3, t^4\} \), where \( t^1, t^2 \) and \( t^3 \) are the same as in Example 3. \( t^4 \) is given below.

\[
\begin{array}{cccc}
x & y & a & b \\
t^4: & 2 & 1 & 2 & 4 \\
\end{array}
\]

The premuneration values and the status quo allocation are both the same as in Example 3. Then \( (\mu, p, t^1, \Pi) \) is blocked* but not blocked.

FACT 1. If \( (\mu, p, t, \Pi) \) is blocked, then \( (\mu, p, t, \Pi) \) is blocked*.

However, condition (11) is too stringent and it is not intuitive for agents to know the payoff evaluated at the true type profile. Intuitively, worst-case desideratum is a conservative way to describe and analyze agents' behavior when information is incomplete. Taking maximum is again a conservative description of agent’ reasoning, which is consistent with the worst-case desideratum.

4.2 Naïve Blocking

This subsection discuss some assumptions that are investigated in the literature. Assumptions 1 and 3 are imposed in LMPS. Assumptions 1-3 are imposed in Bikhchandani (2014, 2017). Bikhchandani (2014) imposed additionally Assumption 4, which is also imposed by Pomatto (2015).

ASSUMPTION 1. (One-Dimensional Type) \( W \subset \mathbb{R} \) and \( F \subset \mathbb{R} \).
**Assumption 2.** (Non-Transferable Utility) No transfer is permitted in the model.

**Assumption 3.** (Increasing and Continuous Utility) The prenumeration functions $\nu_{w,f}$ and $\phi_{w,f}$ are strictly increasing and continuous in $w$ and $f$.

**Assumption 4.** (Knowledge within One Side) It is common knowledge that each worker knows the types of all workers and each firm knows the types of all firms.

A state is said to be naively blocked if improvements of blocking agents can be guaranteed for all possible type profiles specified by the partition profile. Because of Assumption 4, we write type profiles as $(w, f)$ instead of $t$ to distinguish two sides of the market. Because of Assumption 2, market states have no payment components.

**Definition 5.** A state $(\mu, w, f, \Pi)$ is said to be naively blocked by $(i, j)$ if

\[
\nu_{w(i), f(j)} > \nu_{w(i), f(\mu(i))} \text{ for all } (w', f') \in \Pi_i(w, f);
\]

\[
\phi_{w(i), f(j)} > \phi_{w(\mu^{-1}(j)), f(j)} \text{ for all } (w', f') \in \Pi_j(w, f).
\]

Naive blocking is a stronger notion than blocking (Definition 2). However, the following fact shows the implication of Assumptions 1-4: naive blocking is the same as blocking.

**Fact 2.** Under Assumptions 1-4, $(\mu, w, f, \Pi)$ is blocked if and only if $(\mu, w, f, \Pi)$ is naively blocked.

### 4.3 Detection in Blocking

As we know in Subsection 3.2, the blocking pair $(i, j)$ can refine $\Pi_i(t)$ and $\Pi_j(t)$ to evaluating each other. However, it is not obvious that to what extent they can refine $\Pi_i(t)$ and $\Pi_j(t)$. The following self-proving fact implies that they can at most refine $\Pi_i(t)$ and $\Pi_j(t)$ as if they have obtained each other’s information.

**Fact 3.** Fix a state $(\mu, p, t, \Pi)$ and a potential blocking combination $(i, j; p)$. For each $t'$ and each $l = 1, 2, \ldots$, $\Pi_i^{[l]}(t')$ and $\Pi_j^{[l]}(t')$ are measurable w.r.t. $\Pi_i \land \Pi_j$.\(^{11}\)

Since $\Pi_i^{[l]}(t)$ and $\Pi_j^{[l]}(t)$ are measurable w.r.t. $\Pi_i \land \Pi_j$, we know that

\[
\Pi_i^{[l]}(t) \supset \Pi_i \land \Pi_j(t) \quad \text{and} \quad \Pi_j^{[l]}(t) \supset \Pi_i \land \Pi_j(t).
\]

In other words, the consideration sets of $i$ and $j$ is never smaller than the set of type profiles when they aggregate their information.

\(^{11}\)We say that a set is measurable w.r.t. $\Pi_i \land \Pi_j$ if it is measurable w.r.t. the $\sigma$-algebra generated by all unions of sets in $\Pi_i \land \Pi_j$. 
When would \( i \) and \( j \) have consideration sets that are the same as the one given by the meet of their partitions, i.e., \( \Pi_i^{[r]}(t) = \Pi_j^{[r]}(t) = \Pi_i \cap \Pi_j(t) \)? As another polar case, when would \( i \) and \( j \) have consideration sets that are given by their own partitions, i.e., \( \Pi_i^{[r]}(t) = \Pi_i(t) \) and \( \Pi_j^{[r]}(t) = \Pi_j(t) \)? To answer these questions, we reformulate our blocking notion and provide sufficient conditions for polar cases.

We reformulate our blocking notion by introducing an indicator correspondence \( \chi \) that represents for each measurable event the willingness of \( i \) and \( j \) of agreeing on the blocking combination \((i, j; p)\). Intuitively, we use \( Y \) for Yes (may be willing to participate the blocking) and \( N \) for No (never be willing to participate the blocking). First, for each \( \pi \in \Pi_i \cap \Pi_j \),

\[
\chi_i(\pi) := \begin{cases} 
\{Y\} & \text{if } \nu_{i'(t'), i}(j) + p > \nu_{i'(t'), i}(\mu(i)) + p_{i, \mu(i)} \text{ for some } t' \in \pi, \\
\{N\} & \text{otherwise};
\end{cases}
\]

\[
\chi_j(\pi) := \begin{cases} 
\{Y\} & \text{if } \phi_{i'(t'), i}(j) - p > \phi_{i'(t'), i}(\mu(j)) - p_{i, \mu(j)} \text{ for some } t' \in \pi, \\
\{N\} & \text{otherwise}.
\end{cases}
\]

Then for each \( k = i, j \) and each \( \pi \subset T \) that is measurable w.r.t. \( \Pi_i \cap \Pi_j \),

\[
\chi_k(\pi) := \bigcup_{\pi' \in \Pi_i \cap \Pi_j: \pi' \subset \pi} \chi_k(\pi').
\]

Using indicator functions, the iteration of (3)-(4) can be re-written as follows. Define \( \chi_i^{[0]} = \chi_i \), \( \chi_j^{[0]} = \chi_j \), and recursively for \( l = 1, 2, \ldots \) that

1. for each \( \pi \in \Pi_i \cap \Pi_j \),

\[
\chi_i^{[l]}(\pi) := \begin{cases} 
\chi_i^{[l-1]}(\pi) & \text{if } \nu_{i'(t'), i}(j) + p > \nu_{i'(t'), i}(\mu(i)) + p_{i, \mu(i)} \text{ for some } t' \in \pi, \\
\{N\} & \text{otherwise};
\end{cases}
\]

\[
\chi_j^{[l]}(\pi) := \begin{cases} 
\chi_j^{[l-1]}(\pi) & \text{if } \phi_{i'(t'), i}(j) - p > \phi_{i'(t'), i}(\mu(j)) - p_{i, \mu(j)} \text{ for some } t' \in \pi, \\
\{N\} & \text{otherwise}.
\end{cases}
\]

2. for each \( k = i, j \) and each \( \pi \in \Pi_k \),

\[
\chi_k^{[l]}(\pi) := \bigcup_{\pi' \in \Pi_i \cap \Pi_j: \pi' \subset \pi} \chi_k^{[l]}(\pi').
\]

For simplicity, the dependence of \( \chi_i^{[l]} \) and \( \chi_j^{[l]} \) on \((\mu, p, \Pi)\) and \((i, j; p)\) is suppressed in the
notation. Let $l^*$ be a sufficiently large integer such that $\chi_k^{[r+1]} = \chi_k^{[r]}$ for $k = i, j$.

**Remark 2.** In this formulation, blocking explicitly represents the willingness of $i$ and $j$ to rematch. More precisely, a state $(\mu, p, t, \Pi)$ is blocked by $(i, j; p)$ if and only if

$$\chi_i^{[r]}(\Pi_i(t)) = \chi_j^{[r]}(\Pi_j(t)) = \{Y\}.$$  

The following two facts consider two polar cases. Fact 4 provides conditions on $\Pi$ and $\chi$ such that the joint information can be detected from each other’s willingness, i.e., $\Pi_i^{[r]}(t) = \Pi_j^{[r]}(t) = \Pi_i \land \Pi_j(t)$. Fact 5 explores conditions such that no more information than the initial partition can be detected from each other’s willingness, i.e., $\Pi_i^{[r]}(t) = \Pi_i(t)$ and $\Pi_j^{[r]}(t) = \Pi_j(t)$.

**Fact 4.** Fix a state $(\mu, p, t, \Pi)$ and a potential blocking combination $(i, j; p)$. If

1. $\chi_j(\Pi_i(t) \backslash \Pi_i \land \Pi_j(t)) = \chi_i(\Pi_j(t) \backslash \Pi_i \land \Pi_j(t)) = \{N\}$; and
2. $\chi_i(\pi) \neq \chi_j(\pi)$ for all $\pi \in \Pi_i \land \Pi_j$ such that $\pi \subset [\Pi_i \land \Pi_j(t)] \backslash [\Pi_i(t) \cup \Pi_j(t)]$,

then $\Pi_i^{[r]}(t) = \Pi_j^{[r]}(t) = \Pi_i \land \Pi_j(t)$.

**Fact 5.** Fix a state $(\mu, p, t, \Pi)$ and a potential blocking combination $(i, j; p)$. If $\chi_k(\pi) \ni Y$ for all $\pi \in \Pi_k$, $k = i, j$, then

$$\Pi_i^{[r]}(t) = \Pi_i(t) and \Pi_j^{[r]}(t) = \Pi_j(t).$$

## 5 Learning-Blocking Paths

Consider a job market where any pair of firm and worker can freely match themselves, and any agent can freely opt being unmatched. Suppose that agents are myopic, i.e., once an agent or a worker-firm pair finds an opportunity to improve their status quo, they will do so by either standing alone or finding a new partner. In this section, we will fix a realized type assignment function $t^*$ and study the relationship between the set of stable allocations, $\mathcal{S}(t^*)$, and blocking paths formed by a sequence of randomly satisfying blocking combinations (a sequence of rematchings). In particular, we show that with probability one an arbitrary random learning-blocking path converges to an incomplete-information stable state after finitely many rematchings. The analysis of this section extends that of Chen and Hu (2017). To avoid repetition, we omit some discussions.
Given a status quo \((\mu, p, t^*, \Pi)\) where the partition profile is common knowledge, each agent may observe one of the following two situations: (i) there is no rematching; and (ii) there is a rematching that matches \(i\) to \(j\) at salary \(p\).

In Case (i), agents update their information according to the partition \(N^{(\mu,p,\Pi)}\) defined in Subsection 3.3, i.e., the information structure now becomes \(H_{\mu,p}(\Pi)\) defined by (7).

In Case (ii), agents commonly observe the blocking. Additionally, they can distinguish two events: one permits \((i,j;p)\) as a blocking combination and one does not. Denote the partition identifying the blocking combination \((i,j;p)\) by \(B^{(\mu,p,\Pi;i,j;p)}\), i.e., \(B^{(\mu,p,\Pi;i,j;p)}(t') = B^{(\mu,p,\Pi;i,j;p)}(t'')\) if and only if either \((i,j;p)\) blocks both \((\mu,p,t',\Pi)\) and \((\mu,p,t'',\Pi)\) or \((i,j;p)\) is neither a blocking combination for one of them. Moreover, agents commonly know that worker \(i\) and firm \(j\) have observed the type of each other. Denote the partition that indicates the type of an agent \(k\) by \(O^{(k)}\) where \(k \in I \cup J\), i.e., for any \(t', t'' \in T\), \(O^{(k)}(t') = O^{(k)}(t'')\) if and only if \(t'(k) = t''(k)\). Agents update their information according to their observations.

Formally, we say that a state \((\mu', p', t^*, \Pi')\) is derived from another state \((\mu, p, t^*, \Pi)\) by satisfying a blocking combination \((i,j;p)\) for \((\mu, p, t^*, \Pi)\), which is denoted by \((\mu, p, t^*, \Pi)\) \(\xleftarrow{(i,j;p)} (\mu', p', t^*, \Pi')\), if

\[
\begin{align*}
\mu'(i) &= j; & \mu'(\mu^{-1}(j)) &= \emptyset; \\
\mu'(i') &= \mu(i') & \text{for all } i' \in I \text{ s.t. } i' \neq \mu^{-1}(j) \text{ and } i' \neq i; \\
p'_{i,j} &= p; & p'_{(\mu)-1(j), \emptyset} &= 0; \\
p'_{i', \mu'(i')} &= p'_{\mu'(i'), \mu(i')} & \text{for all } i' \in I \text{ s.t. } i' \neq \mu^{-1}(j) \text{ and } i' \neq i,
\end{align*}
\]

and

\[
\begin{align*}
\Pi'_i &= \Pi_i \wedge N^{(\mu,p,\Pi)} \wedge B^{(\mu,p,\Pi;i,j;p)} \wedge O^{(j)}; \quad (12) \\
\Pi'_j &= \Pi_j \wedge N^{(\mu,p,\Pi)} \wedge B^{(\mu,p,\Pi;i,j;p)} \wedge O^{(i)}; \quad (13) \\
\Pi'_{k} &= \Pi_k \wedge N^{(\mu,p,\Pi)} \wedge B^{(\mu,p,\Pi;i,j;p)} \text{ for all } k \neq i, j. \quad (14)
\end{align*}
\]

We formally describe a Learning-Blocking Path below, where \(\mu, p\) and \(\Pi\) are the state variables during the matching process.\(^{12}\) At each stage, \((\mu, p, t^*, \Pi)\) is the status quo.

**Learning-Blocking Path**

**Input.** An arbitrary state \((\mu^0, p^0, t^*, \Pi^0)\).

\(\footnote{\text{A flow chat version is provided in Appendix C.}}\)
**Initialization.** Initialize $\mu$, $p$ and $\Pi$ to be $\mu^0$, $p^0$ and $\Pi^0$, respectively.

**Phase 1.** There are two exclusive cases.

(a) $(\mu, p, t^*, \Pi)$ is blocked by some individual or pair. Go to **Phase 2**.

(b) $(\mu, p, t^*, \Pi)$ is not blocked. Go to **Phase 3**.

**Phase 2.** Derive $(\mu', p', t^*, \Pi')$ such that $(\mu', p', t^*, \Pi') \leftarrow^{(i,j,p)} (\mu, p, t^*, \Pi)$.

Set $(\mu, p, t^*, \Pi)$ to be $(\mu', p', t^*, \Pi')$. Go to **Phase 1**.

**Phase 3.** There are two exclusive cases.

(a) $H_{\mu,p}(\Pi) = \Pi$, i.e. “no rematching” provides no information. Go to **END**.

(b) $H_{\mu,p}(\Pi) \neq \Pi$, i.e. “no rematching” contains information that could help agents refine their partitional information function.

Set $(\mu, p, t^*, \Pi)$ to be $(\mu, p, t^*, H_{\mu,p}(\Pi))$. Go to **Phase 1**.

**END.** Output $(\mu, p, t^*, \Pi)$.

We say that a **LEARNING-BLOCKING PATH** is *finite* if it involves finitely many rematchings, i.e. Phase 2 and Case (b) ii of Phase 3 are triggered finitely many times.

Before stating our results, we need the following assumption.

**Assumption 5.** Payments permitted in the job market are integers.\(^\text{13}\)

Indeed, practically payments are measured in monetary units and hence integers. Obviously, every stable allocation in $\mathcal{S}(t^*)$ could be reached by the **LEARNING-BLOCKING PATH** starting itself.\(^\text{14}\) Given an arbitrary initial state, we show that by carefully choosing blocking pairs to be rematched (if there are many), we can construct a *finite* **LEARNING-BLOCKING PATH** (i.e., a process which consists of finitely many rematchings) that ends with a stable state.

\(^\text{13}\)Salaries must be rounded to the nearest dollar, penny, or mill. This is a technical assumption to ensure *finite* bargaining choices when a worker-firm pair negotiates, as well as, more importantly, a realistic situation in decentralized market practice which we aim to describe. See Crawford and Knoer (1981), Kelso and Crawford (1982), and Chen et al. (2016) for similar integral assumptions when *finite* matching processes are studied. In marriage models where our results hold and there is no payment involved, of course this assumption is not necessary any more.

Moreover, one can easily construct an example, where two firms compete for one worker, the salary increment converges to zero, and the limit salary still permits a blocking. Therefore, finite path cannot be guaranteed even in the complete-information environment without Assumption 5.

\(^\text{14}\)Stable states are well defined under Assumption 5 and the existence is still guaranteed (see (Crawford and Knoer, 1981, Theorem 1)). In the rest of the paper, we refer stability and $\mathcal{S}(t^*)$ as the ones defined under Assumption 5.
Proposition 2. Suppose Assumption 5 holds. Then starting from any arbitrary initial state, there exists a finite Learning-Blocking Path that leads to a stable state.

Now we consider a random process which starts with an arbitrary state, \((\mu^0, p^0, t^*, \Pi^0)\), and then proceed to generate a random Learning-Blocking Path, i.e., whenever an intermediate state is blocked, randomly satisfy a blocking combination if there are many. In the complete-information environment, RV assumes that (in our language) the probability that every particular blocking combination for a blocked allocation \((\mu, p)\) being chosen is bounded away from zero, which stems from a distribution over all complete-information blocking combinations and depends only on the allocation \((\mu, p)\) (and not on the rematching history), i.e., blocking pairs are independently drawn.

With incomplete information, similar restrictions are imposed when we study random Learning-Blocking Paths. In particular, blocking combinations are independently drawn from a distribution over all blocking combinations, where each combination occurs with positive probability and the probability depends only on the state but not on the rematching history. For such distribution on blocking combinations, we obtain the following result which immediately follows from Proposition 2.

**Proposition 2’.** Suppose Assumption 5 holds. Then the random Learning-Blocking Path starting from an arbitrary state converges to a stable state with probability one.

Proposition 2’ predicts the ultimate state of a random matching market, as a corollary of Proposition 2. This random market mimics the real world situations: agents meet and negotiate randomly until they expect no more (utility or profit) improvement.

Although the information updating along a Learning-Blocking Path in the current setting is different from that of Chen and Hu (2017), the proof of Proposition 2 is very similar to that of Theorem 1 in Chen and Hu (2017), which is omitted. Similarly, the robustness discussion in (Chen and Hu, 2017, Subsection 5.1) can be extended to the current setting straightforwardly.

6 Comparative Statics

6.1 New Positions and New Workers

New workers or new positions may cause a market to shift from one to another. Suppose the original market state is stable. With new agents, the stable status may be broken and Learning-Blocking Paths may be triggered. In this section, we investigate the effect of adding agents into a stable market state.
Let $\Gamma = (I, J, t^*, T, \nu, \phi)$ and $\Gamma' = (I', J', t'^*, T', \nu', \phi')$ be two matching markets. We say the market $\Gamma = (I, J, t^*, T, \nu, \phi)$ is consistent with $\Gamma' = (I', J', t'^*, T', \nu', \phi')$ if $\nu' = \nu$, $\phi' = \phi$ and the natural restrictions of $T$ and $T'$ to the set $(I \cap I') \cup (J \cap J')$ coincide, i.e., if for $\bar{I} := I \cap I'$ and $\bar{J} := J \cap J'$ the following two condition holds:

$$
\begin{align*}
& \{ t^*|_{\bar{I} \cup \bar{J}} = t'^*|_{\bar{I} \cup \bar{J}}, \\
& \{ t|_{\bar{I} \cup \bar{J}} \in W^{|I|} \times F^{|J|} : t \in T \} = \{ t|_{\bar{I} \cup \bar{J}} \in W^{|I|} \times F^{|J|} : t \in T' \},
\end{align*}
$$

where $t|_{\bar{I} \cup \bar{J}}$ is the restriction of the vector $t$ to $\bar{I} \cup \bar{J}$. Throughout this section, we take $\Gamma'$ as a one-agent extension of $\Gamma$, i.e., a market that is consistent with $\Gamma$ and has exactly one more worker or one more firm than $\Gamma$.

In the complete information setting, we have the following property if utilities are non-transferable. See Section 2.5 of Roth and Sotomayor (1990) and Blum et al. (1997). The intuition is that expanding one side of the market increases the competition within that side.

**Property.** Adding one worker (firm) to a stable market state, the result of any (learning-) blocking path makes all other workers (firms) weakly worse off and all firms (workers) weakly better off.

However, this property may collapse when information is incomplete. We provide some counterexamples in the next subsection.

### 6.2 Counterexamples

In Examples 5 and 6 below, agent types are correlated, so inference can be made through correlation. In Example 7, while agent types are independent, inference can also be made through colored hat reasoning.\(^\text{15}\)

Throughout this subsection, prenumeration functions are given by

$$
\nu_{wf} = |wf| - 1 \text{ and } \phi_{wf} = 3 + wf. \quad (15)
$$

We assume there is no transfer between workers and firms (regarded as a constant 0).

**Example 5.** *(Adding one worker makes another worker better off.)* Consider a market $\Gamma = (I, J, t^*, T)$ with $I = \{x\}$, $J = \{a\}$, $t^* = (t^*(x), t^*(a)) = (1; 4)$. The set of possible type assignment functions is $T = \{t, t^*\}$, where $t = (-1; 4)$. There are two stable states

\(^{15}\)See LMPS for the colored hat reasoning in the context of matching with incomplete information.
for this market: the one with no match and the one matches worker \( x \) with firm \( a \), i.e., State \( S_1 \) and State \( S_2 \) given in Table 2.\(^{16}\)

<table>
<thead>
<tr>
<th>States</th>
<th>Matches</th>
<th>Partition Profile</th>
<th>Blocking Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_1 )</td>
<td>( \mu(x) = a )</td>
<td>( \Pi_x = \Pi_a = { { t }, { t^* } } )</td>
<td></td>
</tr>
<tr>
<td>( S_2 )</td>
<td>( \mu(x) = \emptyset )</td>
<td>( \Pi_x = { { t }, { t^* } } )</td>
<td>( \Pi_a = { { t, t^* } } )</td>
</tr>
<tr>
<td>( S_3 )</td>
<td>( \mu(x) = \emptyset )</td>
<td>( \Pi_x = \Pi_a = { { t }, { t^* } } )</td>
<td>( (a, x) )</td>
</tr>
</tbody>
</table>

Table 2: Possible states for \( \Gamma \) in Example 5.

Suppose the status quo State \( S_2 \).

Now a new worker enters the market so that we are faced with a new market \( \Gamma' = (I', J', t'^*, T') \), where \( J' = J, I' = \{ x, y \} \), \( t'^* = (t'^*(x), t'^*(y); t'^*(a)) = (1, 1/2, 4) \), and \( T' = \{ t', t^* \} \) where \( t' = (-1, 1; 4) \). It is obvious that \( \Gamma' \) is consistent with \( \Gamma \). For the

<table>
<thead>
<tr>
<th>States</th>
<th>Matches</th>
<th>Partition Profile</th>
<th>Blocking Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_1' )</td>
<td>( \mu(x) = a )</td>
<td>( \Pi_x = \Pi_y = \Pi_a = { { t' }, { t'^* } } )</td>
<td></td>
</tr>
<tr>
<td>( S_2' )</td>
<td>( \mu(x) = \emptyset )</td>
<td>( \Pi_x = \Pi_y = \Pi_a = { { t' }, { t'^* } } )</td>
<td>( (a, x) )</td>
</tr>
<tr>
<td>( S_3' ) or ( S_4' )</td>
<td>( \mu(x) = \emptyset )</td>
<td>( \Pi_a = { { t' }, { t'^* } } ) or ( \Pi_a = { { t', t'^* } } )</td>
<td>( (a, y) )</td>
</tr>
</tbody>
</table>

Table 3: Possible states for \( \Gamma' \) in Example 5.

market \( \Gamma' \), there are only four possible states, listed in Table 3. Among them, the state with a match between \( x \) and \( a \), i.e., State \( S_1' \), is stable. Note that if matched with worker \( y \), firm \( a \) could infer that worker \( x \) has a type of 1. Thus, with this new information, firm \( a \) will be better off if matched with worker \( x \). From \( S_1 \) to a \( S_1' \), worker \( x \) is better off.

Remark 3. The Rural Hospital Theorem (and the Lone Wolf Theorem as a special case, see Roth (1984)) is violated by the market \( \Gamma \). Namely, \( x \) is matched under stable state \( S_1 \) but not matched under stable state \( S_2 \).

Example 6. (Adding one worker makes a firm worse off.) Consider a market \( \Gamma = (I, J, t^*, T) \) with \( I = \{ x \} \), \( J = \{ a, b \} \), \( t^* = (t^*(x); t^*(a), t^*(b)) = (1; 2, 4) \), and \( T = \{ t, t^* \} \) where \( t = (-1; 2, 4) \).

\(^{16}\)We specify the partition profile for each listed states in this example. While to simplify notations, we will state the matching but omit the partition profile whenever the blocking pair, if any, is independent of the partition profile. Moreover, we will list only states that are relevant for discussion, instead of all states.

16
Suppose the status quo stable state is given by \( \mu(x) = a \) and \( \Pi_x = \Pi_a = \{\{t\}, \{t^*\}\} \).

Now a new worker enters the market so that we are faced with a new market \( \Gamma' = (I', J', t^{*'}, T') \), where \( J' = J, \ I' = \{x, y\} \), \( t^{*'} = (t^{*'}(x), t^{*'}(y); t^{*'}(a), t^{*'}(b)) = (1, 1/2; 2, 4) \), and \( T' = \{t', t^{*'}\} \) where \( t' = (-1, 1; 2, 4) \). It is obvious to check that \( \Gamma' \) is consistent with \( \Gamma \).

For the market \( \Gamma' \), there are seven possibilities, listed in the following table.

<table>
<thead>
<tr>
<th>Possibilities</th>
<th>Matches</th>
<th>Blocking Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \mu(x) = \emptyset ) and ( \mu(y) = \emptyset )</td>
<td>( (b, y) ) and ( (a, y) )</td>
</tr>
<tr>
<td>2</td>
<td>( \mu(x) = a ) and ( \mu(y) = \emptyset )</td>
<td>( (b, y) )</td>
</tr>
<tr>
<td>3</td>
<td>( \mu(x) = \emptyset ) and ( \mu(y) = a )</td>
<td>( (b, y) )</td>
</tr>
<tr>
<td>4</td>
<td>( \mu(x) = b ) and ( \mu(y) = \emptyset )</td>
<td>( (a, y) )</td>
</tr>
<tr>
<td>5</td>
<td>( \mu(x) = \emptyset ) and ( \mu(y) = b )</td>
<td>( (a, x) )</td>
</tr>
<tr>
<td>6</td>
<td>( \mu(x) = a ) and ( \mu(y) = b )</td>
<td>( (b, x) )</td>
</tr>
<tr>
<td>7</td>
<td>( \mu(x) = b ) and ( \mu(y) = a )</td>
<td></td>
</tr>
</tbody>
</table>

Note that there is only one stable state, with \( \mu(x) = b, \mu(y) = a \) and \( \Pi_v = \{\{t'\}, \{t^{*'}\}\} \) for all \( v \in I' \cup J' \), which makes firm a worse off.

We have two additional observations pertain to the matching welfare. One is that the number of matched firms increases and the other is that the firm-side aggregate payoff, as well as the social welfare, increases.

**Remark 4.** The Lattice Structure (see Knuth (1976), who attributes the result to John Conway) is violated by \( \Gamma \). Consider two stable matchings \( \mu(x) = a \) and \( \mu(x) = b \). The firm join results in \( \mu^{-1}(a) = x = \mu^{-1}(b) \), which is not a proper matching.

**Example 7.** (Adding one firm makes a firm better off and a worker worse off.) Consider a market \( \Gamma = (I, J, t^*, T) \) with \( I = \{x, y\}, J = \{a\}, \) and \( T = \{t_1, t_2, t_3, t^*\} \) where

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>t_1</td>
<td>-1</td>
<td>-4</td>
</tr>
<tr>
<td>t_2</td>
<td>1</td>
<td>-4</td>
</tr>
<tr>
<td>t_3</td>
<td>-1</td>
<td>3/2</td>
</tr>
<tr>
<td>t^*</td>
<td>1</td>
<td>3/2</td>
</tr>
</tbody>
</table>

Suppose the status quo stable state is given by \( \mu(x) = a, \Pi_x = \{\{t_1, t_3\}, \{t_2, t^*\}\}, \Pi_y = \{\{t_1, t_3\}, \{t_2, t^*\}\} \) and \( \Pi_a = \{\{t_1, t_3\}, \{t_2, t^*\}\} \).

Now a new firm enters the market so that we are faced with a new market \( \Gamma' = \).
\[(I', J', t^s', T')\), where \(J' = \{a, b\}\), \(I' = I\), and \(T = \{t^{1'}, t^{2'}, t^{3'}, t^{s'}\}\)

where

\[
\begin{array}{cccc}
  x & y & a & b \\
  t^{1'} & -1 & -4 & 4 & 1/2 \\
  t^{2'} & 1 & -4 & 4 & 1/2 \\
  t^{3'} & -1 & 3/2 & 4 & 1/2 \\
  t^{s'} & 1 & 3/2 & 4 & 1/2 \\
\end{array}
\]

It is obvious that \(\Gamma'\) is consistent with \(\Gamma\).

<table>
<thead>
<tr>
<th>Possibilities</th>
<th>States (Partial Description)</th>
<th>Blocking Pairs/Updating/IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(\mu(x) = \emptyset) and (\Pi_a = {{t^{1'}, t^{2'}, t^{3'}, t^{s'}}})</td>
<td>(\Pi_a = {{t^{1'}, t^{2'}, t^{3'}, t^{s'}}})</td>
</tr>
<tr>
<td>2</td>
<td>(\mu(x) = \emptyset) and (\Pi_a = {{t^{1'}, t^{2'}, t^{3'}, t^{s'}}})</td>
<td>(\Pi_a = {{t^{1'}, t^{2'}, t^{3'}, t^{s'}}})</td>
</tr>
<tr>
<td>3</td>
<td>(\mu(x) = a, \mu(y) = \emptyset) and (\Pi_a = {{t^{1'}, t^{3'}, t^{2'}, t^{s'}}})</td>
<td>(\Pi_a = {{t^{1'}, t^{2'}, t^{3'}, t^{s'}}})</td>
</tr>
<tr>
<td>4</td>
<td>(\mu(x) = a, \mu(y) = \emptyset) and (\Pi_a = {{t^{1'}, t^{3'}, t^{2'}, t^{s'}}})</td>
<td>(\Pi_a = {{t^{1'}, t^{2'}, t^{3'}, t^{s'}}})</td>
</tr>
<tr>
<td>5</td>
<td>(\mu(x) = \emptyset) and (\mu(y) = a)</td>
<td>(\Pi_a = {{t^{1'}, t^{2'}, t^{3'}, t^{s'}}})</td>
</tr>
<tr>
<td>6</td>
<td>(\mu(x) = b \text{ and } \mu(y) = \emptyset)</td>
<td>Not IR</td>
</tr>
<tr>
<td>7</td>
<td>(\mu(x) = \emptyset \text{ and } \mu(y) = b)</td>
<td>Not IR</td>
</tr>
<tr>
<td>8</td>
<td>(\mu(x) = b \text{ and } \mu(y) = a)</td>
<td>Not IR</td>
</tr>
</tbody>
</table>

Note that there is only one stable matching with \(\mu(x) = \emptyset\) and \(\mu(y) = a\), regardless of the partition profile, which makes firm \(a\) better off and at the same time makes worker \(x\) worse off.

**Remark 5.** Even if we allow for transfers, it cannot be used by firms to screen worker types in our examples. Particularly, whenever a firm wants to screen out a low type worker by setting a salary that he would never accept, the high type worker always finds the salary unacceptable as well. So the violation of Property cannot be eased by introducing transfer.

### 6.3 Restoring Comparative Statics

A common feature of the examples in the last subsection, in addition to zero payments, is that all status quo states are stable, while none of the allocations is complete-information stable. Thus a natural question to ask is whether we can restore Property for the class of markets where utilities are non-transferable and the status quo allocation is complete-information stable. For this class of markets, the following proposition restores
\textbf{PROPERTY} under the mild assumption of strict preferences, i.e., \(\phi_{t^i(i),t^j(j)} \neq \phi_{t^{i'}(i'),t^{j'}(j)}\) for \(i \neq i'\) and \(\nu_{t^i(i),t^j(j)} \neq \nu_{t^{i'}(i'),t^{j'}(j')}\) for \(j \neq j'\). Recall that \(\Gamma'\) is a one-agent extension of \(\Gamma\).

\textbf{PROPOSITION 3.} Suppose preferences are strict and Assumption 2 holds. If \((\mu, t^*, \Pi)\) is a stable \(\Gamma\)-state such that \(\mu\) is a complete-information stable allocation, then for any stable \(\Gamma'\)-state \((\mu', t^{*'}, \Pi')\) produced by Learning-Blocking Paths, the payoffs of all existing firms (resp. workers) increases and the payoffs of all existing firms (resp. workers) decreases when \(J \subset J'\) (resp. \(I \subset I'\)), compared with the payoffs under \((\mu, t^*, \Pi)\).

The welfare comparison is between a complete-information stable \(\Gamma\)-allocation and an incomplete-information stable \(\Gamma'\)-allocation, the latter may not be complete-information stable as the following example shows.

\textbf{EXAMPLE 8.} Consider a market \(\Gamma = (I, J, t^*, T)\) with \(I = \{x, y\}\), \(J = \{a\}\), \(t^* = (t^*(x), t^*(y), t^*(a)) = (5, 4; 1)\). The set of possible type assignment functions is \(T = \{t, t^*\}\), where \(t = (5, -4, 1)\). Premuneration functions are given by (15).

Suppose the status quo is given by \(\mu(x) = a, \mu(y) = \emptyset, \Pi_y = \{\{t\}, \{t^*\}\}\) and \(\Pi_a = \Pi_x = \{\{t, t^*\}\}\).

Now a new firm enters the market so that we are faced with a new market \(\Gamma' = (I', J', t'^*, T')\), where \(J' = \{a, b\}\), \(I' = I\), \(t'^* = (t'^*(x), t'^*(y); t'^*(a), t'^*(b)) = (5, 4; 1, 2)\), and \(T' = \{t', t'^*\}\) where \(t' = (5, -4; 1, 2)\). It is obvious to check that \(\Gamma'\) is consistent with \(\Gamma\).

We find that \((x, b)\) is a blocking pair for the stable \(\Gamma\)-status quo in \(\Gamma'\). Satisfying \((x, b)\) results in a stable \(\Gamma'\)-state that is given by \(\mu(x) = b, \mu(y) = \emptyset, \Pi_y = \{\{t\}, \{t^*\}\}\) and \(\Pi_a = \Pi_b = \Pi_x = \{\{t, t^*\}\}\), where the allocation is not complete-information stable in \(\Gamma'\).

Assumption 2 is imposed because of the following example. It says that even in the complete-information environment, \textbf{PROPERTY} cannot be guaranteed if utility is transferable.

\textbf{EXAMPLE 9.} Consider a market \(\Gamma = (I, J, t^*, T)\) with \(I = \{x, y\}\), \(J = \{a\}\), \(t^* = (t^*(x), t^*(y), t^*(a)) = (1, 2; 3)\). The set of possible type assignment functions is \(T = \{t^*\}\), i.e., information is complete. Premuneration functions are given by \(\nu_{a,f} = \phi_{a,f} = w_f\).

Suppose the status quo is given by \(\mu(y) = a, \mu(x) = \emptyset, \) and \(p_{y,a} = 0\).

Now a new firm enters the market so that we are faced with a new market \(\Gamma' = (I', J', t'^*, T')\), where \(J' = \{a, b\}\), \(I' = I\), \(t'^* = (t'^*(x), t'^*(y); t'^*(a), t'^*(b), t'^*(c)) = (1, 2; 3, 2)\), and \(T' = \{t'^*\}\), i.e., information is again complete. It is obvious to check that \(\Gamma'\) is consistent with \(\Gamma\).
We explicitly express a Learning-Blocking Path in Table 6, which shows that the payoff of worker $y$ decreases in the new stable state.

<table>
<thead>
<tr>
<th>States</th>
<th>Matches</th>
<th>Payoffs</th>
<th>Blocking Combinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>$\mu(y) = a$</td>
<td>$p_{y,a} = 0$</td>
<td>$(y,b;3)$</td>
</tr>
<tr>
<td></td>
<td>$\mu(x) = \emptyset$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_2$</td>
<td>$\mu(y) = b$</td>
<td>$p_{y,b} = 3$</td>
<td>$(x,b;-1)$</td>
</tr>
<tr>
<td></td>
<td>$\mu(x) = \emptyset$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_3$</td>
<td>$\mu(x) = b$</td>
<td>$p_{x,b} = -1$</td>
<td>$(y,a;-1)$</td>
</tr>
<tr>
<td></td>
<td>$\mu(y) = \emptyset$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_4$</td>
<td>$\mu(x) = b$</td>
<td>$p_{x,b} = -1$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu(y) = a$</td>
<td>$p_{y,a} = -1$</td>
<td></td>
</tr>
</tbody>
</table>

7 Ex Ante Stability

7.1 Stability

Our paper is not the first to introduce a stability notion for matching with incomplete information. Particularly, our notion is closely related to that of LMPS. The notion of LMPS is *ex ante* in that it is independent of the true type assignment function. One can imagine that there is an outside analyst who knows the model but not anything about agents’ partitional information, and wants to predict possible outcomes for the market. The notion of blocking is designed to only exclude outcomes that they can certainly be “blocked.” Formally, an (ex post) outcome $(\mu, p, t)$ specifies an allocation and a type assignment function. The individual rationality of an ex post outcome is defined in the same manner as Definition 1. Let $\Sigma$ be a set of ex post outcomes.

Denote the binary partition that indicates whether $(\mu, p, t') \in \Sigma$ or not by $T_{\Sigma,\mu,p}$, i.e., for any $t', t'' \in T$, $T_{\Sigma,\mu,p}(t') = T_{\Sigma,\mu,p}(t'')$ if and only if either both $(\mu, p, t') \in \Sigma$ and $(\mu, p, t'') \in \Sigma$ or neither of them. Recall that $O^{(k)}$ is the partition that indicates the type of an agent $k$ where $k \in I \cup J$, i.e., for any $t', t'' \in T$, $O^{(k)}(t'') = O^{(k)}(t''')$ if and only if $t'(k) = t''(k)$. Now we define the partition profile $\Pi_{\Sigma,\mu,p}$ as follows:

$$\Pi_{\Sigma,\mu,p,i} := O^{(i)} \land T_{\Sigma,\mu,p} \text{ for all } i \in I;$$

$$\Pi_{\Sigma,\mu,p,j} := O^{(j)} \land T_{\Sigma,\mu,p} \text{ for all } j \in J.$$
Definition 6. Fix a nonempty set of individually rational outcomes \( \Sigma \). A matching outcome \((\mu, p, t) \in \Sigma \) is said to be **\( \Sigma \)-blocked** if \((\mu, p, t, \Pi_{\Sigma, \mu, p}) \) is blocked. A matching outcome \((\mu, p, t) \in \Sigma \) is **\( \Sigma \)-stable** if it is not \( \Sigma \)-blocked.

The set of outcomes that are immune to the blocking described above is given by the iteration process below. Let \( \Sigma^0 \) be the set of all individually rational outcomes. For \( l \geq 1 \), define
\[
\Sigma^l := \{ (\mu, p, t) \in \Sigma^{l-1} : (\mu, p, t) \text{ is } \Sigma^{l-1} \text{-stable} \}.
\]

The set of *ex ante stable outcomes* is given by
\[
\Sigma^\infty := \bigcap_{l=1}^{\infty} \Sigma^l.
\]

If \((\mu, p, t)\) is an ex ante stable outcome, we say that the allocation \((\mu, p)\) is an ex ante stable allocation at \( t \).

7.2 Fixed Point Characterization

The iterative procedure of (18) describes an algorithm to obtain the set of all ex ante stable allocations. This set has a fixed point characterization, which is often more convenient for verifying that a given matching outcome is ex ante stable. The analysis of this subsection extends that of Subsection 3.4 of LMPS.

Definition 7. A nonempty set of individually rational matching outcomes \( E \) is **self-stabilizing** if every \((\mu, p, t) \in E \) is \( E \)-stable. The set \( E \) **stabilizes a given matching outcome** \((\mu, p, t) \in E \) and \( E \) is self-stabilizing. A set of type assignments \( T^* \subset T \) **stabilizes** an allocation \((\mu, p)\) if \( \{ (\mu, p, t) : t \in T^* \} \) is a self-stabilizing set.

The following proposition provides a fixed-point characterization of the set of ex ante stable outcomes.

Proposition 4. The following statements are true.

1. If \( E \) is a self-stabilizing set, then \( E \subset \Sigma^\infty \).
2. \( \Sigma^\infty \) is a self-stabilizing set and hence the largest self-stabilizing set.
7.3 Equivalence and Connections with Stability

The following proposition shows the connection between stable states and stable outcomes in $\Sigma^\infty$. It extends Theorem 2 in Chen and Hu (2017).\(^{17}\)

**Proposition 5.** $(\mu, p, t) \in \Sigma^\infty$ if and only if there exists a partition profile $\Pi$ such that $(\mu, p, t, \Pi)$ is stable.

Recall that $\mathcal{S}$ is the ex ante set of stable allocations defined in Subsection 3.4. The following corollary is immediate from Proposition 5.

**Corollary 1.** $\mathcal{S} = \{ (\mu, p) \in \mathcal{A} : \exists t \in T \text{ s.t. } (\mu, p, t) \in \Sigma^\infty \}$.\(^{29}\)

The set of stable allocations actually depends on the ground set $T$. This dependence is indicated by the notation $\mathcal{S}$. The following corollary says that the set of stable allocations is increasing in the ground set of possible type assignment functions.

**Corollary 2.** If $T \subset T'$, then $\mathcal{S}(T) \subset \mathcal{S}(T')$.

**Remark 6.** The notion of stability (Definition 3) coincides with the notion ex ante stability when the only source of firms’ heterogeneous information is that each firm can observe the type of her own employee. To make a precise comparison, consider a stable outcome $(\mu, p, t) \in \Sigma^\infty$. Let $\Pi^0$ be the partition profile over $T$ that distinguishes for each agent only the type of herself/himself and her/his partner (if any). We consider the iteration starting with $\Pi^0$ and for $l = 1, 2, \ldots$, $\Pi^l := H_{\mu,p}(\Pi^{l-1})$. The intuition is that $(\mu, p, t, \Pi^\infty)$ must be a stable state in the sense of Definition 3.

Since all complete-information stable outcomes are included in $\Sigma^\infty$, we can compare the above iteration with the one in Remark 3 for those outcomes. A main difference between two iterations, and thus two stability notions, is that we allow for arbitrary starting partition profile $\Pi^0$ in Remark 3. For outcomes that are in $\Sigma^\infty$ but not complete-information stable, stability (Definition 3) still allows for flexible $\Pi$ to make $(\mu, p, t, \Pi)$ stable, instead of a particular $\Pi^\infty$.

\(^{17}\)As a special case, Chen and Hu (2017) discussed the stability notion under the following assumption. For any two type assignment functions, the firm components of them are the same, which is commonly known and denoted by $f$, i.e.,

**Assumption.** For any $t', t'' \in T$ where $t' = (w', f')$ and $t'' = (w'', f'')$, $f' = f'' = f$.

Unlike the case for two-sided incomplete information, we do not need to impose the non-empty conditions as in Definition 2.
8 Bayesian Stability

8.1 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.” (Bikhchandani (2017)) Assume it is common knowledge that the type assignment function follows the cumulative probability distribution function $G$, where $\text{supp}(G) = T$. For a subset $T'$ of $T$, we use the notation $G(\cdot|T')$ to denote the distribution of type assignment functions restricted to $T'$. The notation $\tilde{t}$ is for a random variable drawn from $G$ or $G(\cdot|T')$.

Recall that a state of the matching market is represented by $(\mu, p, t, \Pi)$. Individual rationality is defined as in Section 3. For the given $\Pi$ and $(i,j;p)$, we define the indicator function $\chi$ over $\Pi_i \land \Pi_j$ that represents agents’ willingness of participating the blocking. Particularly, the value is yes if and only if there is expected improvement, i.e., for each $\pi \in \Pi_i \land \Pi_j$,

$$\chi_i(\pi) := \begin{cases} \{Y\} & \text{if } \mathbb{E}[\nu_{t(i),t(j)}|\pi] + p > \nu_{t(i),t(\mu(i))} + p_{t,\mu(i)} \\ \{N\} & \text{otherwise} \end{cases}$$

$$\chi_j(\pi) := \begin{cases} \{Y\} & \text{if } \mathbb{E}[\phi_{t(i),t(j)}|\pi] - p > \phi_{t(\mu^{-1}(j)),t(j)} - p_{\mu^{-1}(j),j} \\ \{N\} & \text{otherwise} \end{cases}$$

for each $k = i, j$ and each $\pi \in \Pi_k$,

$$\chi_k(\pi) := \bigcup_{\pi' \in \Pi_i \land \Pi_j : \pi' \subset \pi} \chi_k(\pi').$$

Similar to (3)-(4), the potential blocking pair $(i, j)$ conservatively refine their consideration set. Intuitively, an agent considers $t''$ as a relevant type profile only if her/his potential partner may have an expected improvement. Define $\Pi_0[0] = \Pi$ and recursively for $l = 1, 2, \ldots$ that

$$\Pi_i[l](t') := \left\{ t'' \in \Pi_i(t') : Y \in \chi_j(\Pi_{j}[l-1](t'')) \right\}$$

$$\Pi_j[l](t') := \left\{ t'' \in \Pi_j(t') : Y \in \chi_i(\Pi_{i}[l-1](t'')) \right\},$$

where the dependence of $\Pi_i[l]$ and $\Pi_j[l]$ on $(\mu, p, \Pi)$ and $(i, j; p)$ is suppressed in the notation.
**Definition 8.** A state \((\mu, p, t, \Pi)\) is said to be **Bayesian blocked** if there exists a worker-firm pair \((i, j)\) and a payment \(p \in \mathbb{R}\) such that \(\Pi^l_i(t) \neq \emptyset\), \(\Pi^l_j(t) \neq \emptyset\) and

\[
\mathbb{E} \left[ \nu_{t(i), t(j)} | \Pi^l_i(t) \right] + p > \nu_{t(i), t(\mu(i))} + p_{t(i)} \quad \text{and} \quad \mathbb{E} \left[ \phi_{t(i), t(j)} | \Pi^l_j(t) \right] - p > \phi_{t(\mu^{-1}(j)), t(j)} - p_{\mu^{-1}(j), j}.
\] (21)

Similar to Lemma 2, we formalize the intuition that more precise information makes it easier to form a blocking, whose proof is the same as that of Lemma 2 and thus omitted.

**Lemma 3.** Suppose \((\mu, p, t, \Pi')\) and \((\mu, p, t, \Pi)\) are two states such that \(\Pi'\) is coarser than \(\Pi\). If \((\mu, p, t, \Pi')\) is Bayesian blocked, then \((\mu, p, t, \Pi)\) is Bayesian blocked.

The Bayesian binary partition \(N(\mu, p, \Pi)\) and Bayesian operator \(H_{\mu, p}\) are defined in the same manner as their counterparts in Subsection 3.3.

**Definition 9.** A state \((\mu, p, t, \Pi)\) is said to be **Bayesian stable** if

1. it is individually rational,
2. it is not Bayesian blocked, and
3. \(\Pi\) is a fixed point of \(H_{\mu, p}\), i.e. \(H_{\mu, p}(\Pi) = \Pi\).

A straightforward observation is that the existence of a Bayesian stable state is guaranteed by the existence of a complete-information stable state.

We say that a state \((\mu, p, t, \Pi)\) is **Bayesian quasi-stable** if \((\mu, p, t, H^k_{\mu, p}(\Pi))\) is Bayesian stable for some \(k \in \mathbb{N}\). We conclude this subsection by an observation that analogous to Proposition 1, whose proof is omitted.

**Proposition 6.** Suppose \((\mu, p, t, \Pi')\) and \((\mu, p, t, \Pi)\) are two states such that \(\Pi'\) is coarser than \(\Pi\). If \((\mu, p, t, \Pi)\) is Bayesian stable, then \((\mu, p, t, \Pi')\) is Bayesian quasi-stable.

### 8.2 Bayesian Stability with One-sided Incomplete Information: Bikhchandani (2017)

The discussion in the last subsection includes the environment with one-sided incomplete information as a special case. Particularly, when the firm-side information is public knowledge,

\[
\mathbb{E} \left[ \nu_{t(i), t(j)} | \Pi^{[l]}_i(t) \right] = \nu_{t(i), t(j)} \quad \text{and} \quad \mathbb{E} \left[ \phi_{t(i), t(j)} | \Pi^{[l]}_j(t) \right] = \mathbb{E} \left[ \phi_{t(i), t(j)} | \Pi^{[1]}_j(t) \right].
\]
Therefore, a state \((\mu, p, t, \Pi)\) is Bayesian blocked if there exists a worker-firm pair \((i, j)\) and a payment \(p \in \mathbb{R}\) such that

\[
\nu_{t(i), t(j)} + p > \nu_{t(i), t(\mu(i))} + p_{t, \mu(i)} \quad \text{and} \quad \mathbb{E} \left[ \phi_{t(i), t(j)} | \Pi^{[1]}_{t(i), t(j)}(t) \right] - p > \phi_{t(\mu^{-1}(j)), t(j)} - p_{\mu^{-1}(j), j}.
\]

The definitions of Bayesian \(N^{(\mu, p, \Pi)}\) and Bayesian stability is the same as in the last subsection. Given a type assignment function \(t\), we denote by \(S^{B}_B(t)\) the set of allocations that can arise in Bayesian stable states and by \(S^{B}\) the ex ante set of Bayesian stable allocations.

In the rest of this subsection, we compare our Bayesian stability notion with the notion studied in Bikhchandani (2017).\(^{18}\) Using our notations, an individually rational matching outcome \((\mu, t) \in \Gamma\) is Bayesian \(\Gamma\)-stable if it is not Bayesian \(\Gamma\)-blocked. A set \(\Gamma\) is Bayesian self-stabilizing if every \((\mu, t) \in \Gamma\) is Bayesian \(\Gamma\)-stable. Let \(\Gamma^0\) be the set of individually rational matching outcomes. For \(l \geq 1\), define

\[
\Gamma^l := \{ (\mu, t) \in \Gamma^{l-1} : (\mu, t) \text{ is Bayesian } \Gamma^{l-1} - \text{stable} \}.
\]

For each matching \(\mu\), the probability distribution over a worker’s type is updated at each stage \(l\) by eliminating types that, together with \(\mu\), would be Bayesian \(l - 1\)-blocked. The set of Bayesian stable matching outcomes is

\[
\Gamma^\infty := \bigcap_{l=0}^{\infty} \Gamma^l.
\]

Let \(\Gamma^\infty(t)\) be the Bayesian stable matchings that are associated with \(t\).

**Proposition 7.** \(\Gamma^\infty(t) \subset S^{B}(t)\).

Example 2 in Bikhchandani (2017) further shows that \(\Gamma^*(t) \subsetneq S^{B}(t)\). Thus, the Bayesian stability notion in this paper includes that of Bikhchandani (2017) in the allocation sense.

Note that \(\Gamma^*(t)\) is not always nonempty, as shown in Proposition 3 and Example 2 of Bikhchandani (2017). If \(\Gamma^*(t) = \emptyset\), then Proposition 7 trivially holds. To prove the proposition, we first argue that for iteration \((23)\), \(\Gamma^{[\ell]+1} = \Gamma^{[\ell]}\), i.e., the sequence \(\{\Gamma^k\}_k\) converges in finite time. This can be shown using the argument we prove Lemma 4, except

\(^{18}\)We make two assumptions so that two notions are discussed under the same environment: anonymous preferences and non-transferable utility. See (Bikhchandani, 2017, pp. 377).
for the difference in notations. \( \Gamma^\infty \) is Bayesian self-stabilizing because \( \Gamma^\infty = \Gamma[^T + 1] = \Gamma[^T] \). The rest of the proof is similar to the necessity part of Proposition 5 and thus omitted.

9 Concluding Remarks

In the previous sections we assumed that agents can observe the type of their partners, which follows from LMPS and CH. This assumption makes the analysis simple but it appears stringent. To see this, let the type of a worker has two components: the worker’s characteristics, such as productivity, and his preference. A firm who hires this worker may observe his productivity but not necessarily his preference. Moreover, different production procedures may reveal different aspects of the worker’s characteristics, i.e., firms may observe the worker’s productivity for one task but not necessarily for others.

Nevertheless, our approach applies to models with more general agent types and more general assumptions about observability, which incorporate the following features. First, an agent’s type is partially observable to her/his matched partner. Second, different components of an agent’s type is observable if matched to different partners. Next, in any state the partition profile is consistent with agents’ observation. Finally, an agent’s premuneration value is determined by the type components that she/he has observed.

More formally, we discuss general agent types and observability by adjusting the model setup in Section 2 as below. The type of each worker is described by \(|J| + 1\) components, i.e., for each \(i\) and each \(t\), \(t(i) = (\tau_0(i), (\tau_j(i))_{j \in J})\), where \(\tau_0(i)\) contains all information about worker \(i\) and \(\tau_j(i)\) represents the partial information about worker \(i\) that is observable to firm \(j\) if they are matched with each other. Similarly, the type of each firm is described by \(|I| + 1\) components, i.e., for each \(j\) and each \(t\), \(t(j) = (\tau_0(j), (\tau_i(j))_{i \in I})\). A state of the matching market, \((\mu, p, t, \Pi)\), specifies an allocation \((\mu, p)\), a type assignment function \(t\) and a partition profile \(\Pi\) such that each agent’s partition is consistent with her/his own type and a component of her/his partner’s type observed by her/him, i.e., for every \(i \in I\) and every \(t'\), \(t''(i) = t'(i)\) and \(\tau''_0(\mu(i)) = \tau'_0(\mu(i))\) for all \(t'' \in \Pi_i(t')\); for every \(j \in J\) and every \(t'\), \(t''(j) = t'(j)\) and \(\tau''_j(\mu^{-1}(j)) = \tau'_j(\mu^{-1}(j))\) for all \(t'' \in \Pi_j(t')\). Finally, we restrict our attention to premuneration functions such that for each \(t\), each \(i\) and each \(j\), \(\nu_{t(i), t(j)}\) depends on \(t(j)\) only through \(\tau_i(j)\); similarly \(\phi_{t(i), t(j)}\) depends on \(t(i)\) only through \(\tau_j(i)\). The generalized formulation include our setup in Section 2 as a special case if for each \(i\), \(\tau_j(i) = \tau_0(i)\) for all \(j\) and for each \(j\), \(\tau_i(j) = \tau_0(j)\) for all \(i\). Furthermore, it is straightforward (but tedious) to replicate our analysis to the setting with general types and assumptions about observability.

Moreover, our approach applies to models with general preferences, where neither
premuneration value functions and transfers are involved. Consider a model where an agent’s type consists of two components: characteristics and preference. Workers’ preferences are ordered list of firm characteristics and vice versa. This flexible formulation includes premuneration value functions as a special case when agents have common preference over the characteristics of agents on the other side of the market. Again, it is straightforward (but tedious) to replicate our analysis to the setting with general preferences.

References


**APPENDICES**

**A PROOFS FOR SECTION 3**

*Proof of Lemma 1.* Suppose to the contrary that

\[ \nu_{t(i),t(j)} + p \leq \nu_{t(i),t(\mu(i))} + p_{t,\mu(i)}. \]

Then (5) implies that \( t \notin \Pi^{[r]}_i(t) \), which by (3) implies either that \( \Pi^{[r]}_j(t) = \emptyset \) or that

\[
\max_{i \in \Pi^{[r]}_j(t)} [\phi_{t(i),t(j)} - p] - [\phi_{t(\mu^{-1}(j)),t(j)} - p_{\mu^{-1}(j),j}] \leq 0. \tag{24}
\]

Since \( (\mu, p, t, \Pi) \) is blocked, we have \( \Pi^{[r]}_j(t) \neq \emptyset \). Moreover, (24) contradicts (6). Therefore, the hypothesis is not true.

The argument for \( \phi_{t(i),t(j)} - p > \phi_{t(\mu^{-1}(j)),t(j)} - p_{\mu^{-1}(j),j} \) is symmetric. \qed
\textbf{Proof of Lemma 2.} Denote the blocking combination for \((\mu, p, t, \Pi')\) by \((i, j; p)\). Since \(\Pi'\) is coarser than \(\Pi\), we know that for each \(t' \in T\),

\[\Pi_i(t') \subset \Pi_i'[t'] \text{ and } \Pi_j(t') \subset \Pi_j'[t'].\]

Note that \(\Pi_i[t']\) defined in (3) is increasing in both \(\Pi_i(t')\) and \(\Pi_i'[t']\), and that \(\Pi_j[t']\) defined in (4) is increasing in both \(\Pi_j(t')\) and \(\Pi_j'[t']\). It follows from induction that for each \(t' \in T\) and each \(l = 1, 2, \ldots\),

\[\Pi_i[l](t') \subset \Pi_i[l](t')\] and \(\Pi_j[l](t') \subset \Pi_j[l](t').\]

Particularly, we have that for each \(t' \in T\),

\[\Pi_i[l]t'(t') \subset \Pi_i[l]t'(t')\] and \(\Pi_j[l]t'(t') \subset \Pi_j[l]t'(t'),\]

where the dependence of \(l^*\) on \(\Pi'\) and \(\Pi\) is suppressed in the notation. The statement follows from Definition 2. \hfill \Box

\textbf{Proof of Proposition 1.} Obviously, the individual rationality of \((\mu, p, t, \Pi)\) is equivalent to the individual rationality of \((\mu, p, t, \Pi')\). Since \((\mu, p, t, \Pi)\) is stable and thus not blocked, we know by Lemma 2 that \((\mu, p, t, \Pi')\) is not blocked. Similarly, \((\mu, p, t', \Pi')\) is not blocked for all \(t' \in N^{(\mu, p, \Pi)}(t)\). (Otherwise, \((\mu, p, t', \Pi)\) is blocked, contradicting \(t' \in N^{(\mu, p, \Pi)}(t)\).) Therefore, \(N^{(\mu, p, \Pi)}(t) \subset N^{(\mu, p, \Pi')}(t)\). As a result, \(|H_{\mu,p}(\Pi')|k(t') = \Pi_k(t')\) for all \(t' \in N^{(\mu, p, \Pi)}(t)\) and all \(k \in I \cup J\).

Since \((\mu, p, t', \Pi')\) is not blocked for all \(t' \in N^{(\mu, p, \Pi)}(t)\), we know that \((\mu, p, t', H_{\mu,p}(\Pi'))\) is not blocked for all \(t' \in N^{(\mu, p, \Pi)}(t)\), which implies that \(N^{(\mu, p, \Pi)}(t) \subset N^{(\mu, p, H_{\mu,p}(\Pi'))}(t)\). Keep applying this argument and define \(\Pi^{k} := H_{\mu,p}(\Pi^{k-1})\) for \(k = 1, 2, \ldots\), until we find a fixed point of \(H_{\mu,p}\). This must be done within finitely many times because the partition profile gets finer whenever it is not a fixed point. Denote the fixed point by \(\Pi^\infty\). Then induction shows that \(N^{(\mu, p, \Pi)}(t) \subset N^{(\mu, p, \Pi^\infty)}(t)\) and that \((\mu, p, t', \Pi^\infty)\) is not blocked for all \(t' \in N^{(\mu, p, \Pi)}(t)\), particularly for \(t' = t\). \hfill \Box

\textbf{B \ PROOFS FOR SECTION 4}

\textbf{Proof of Fact 1.} We have already know that \(\Pi_k[1](t') \subset \Pi_k[1](t')\) for all \(k \in I \cup J\) and \(t' \in T\). By induction, for all \(l \in \mathbb{N}\) (particularly \(l^*\)), we know that \(\Pi_k[l](t') \subset \Pi_k[l](t')\) for all \(k \in I \cup J\) and \(t' \in T\). Since \(\Pi_k[1](t') = \Pi_k[1](t')\) for all \(l = 2\), we have \(\Pi_k[l](t') \subset \Pi_k[l](t')\) for all \(k \in I \cup J\) and \(t' \in T\). The statement follows. \hfill \Box
Proof of Fact 2. The sufficiency is trivial. Note that if \((\mu, w, f, \Pi)\) is blocked, then Lemma 1 implies that
\[
\nu_{w(i), f(j)} > \nu_{w(i, f(\mu(i)))} \quad \text{and} \quad \phi_{w(i), f(j)} > \phi_{w(i, f(\mu(i)))}.
\]
By Assumption 3, the first inequality implies that \(f(j) > f(\mu(i))\). For each \((w', f') \in \Pi_j(w, f)\), Assumption 4 implies that \(f' = f\). Thus, for each \((w', f') \in \Pi_j(w, f)\), we have \(\nu_{w(i), f(j)} - \nu_{w(i), f'(\mu(i))} > 0\) by Assumption 3 again. Since \((w', f') \in \Pi_i(w', f')\), we know that
\[
\max_{(w', f') \in \Pi_i(w', f')} \nu_{w(i), f(j)} - \nu_{w(i), f'(\mu(i))} > 0.
\]
Therefore, \(\Pi^{[1]}_j(w, f) = \Pi_j(w, f)\). By similar argument, \(\Pi^{[1]}_i(w, f) = \Pi_i(w, f)\). Thus, \(\Pi^*_j(w, f) = \Pi_j(w, f)\) and \(\Pi^*_i(w, f) = \Pi_i(w, f)\). The statement follows immediately.  

C A Learning-Blocking Path

```
Is \((\mu, p, t^*, \Pi)\) blocked?
```

arb.

```
Arbitrarily choose a blocking combination \((i, j; p)\).
\((\mu', p', t^*, \Pi') \leftarrow (i, j; p) \quad (\mu, p, t^*, \Pi).
Set \((\mu, p, t^*, \Pi) = (\mu', p', t^*, \Pi').
```

```
Is \(H_{\mu, p}(\Pi) = \Pi?\)
```

```
Y
```

```
\((\mu, p, t^*, \Pi) := (\mu^0, p^0, t^*, \Pi^0).
```

```
N
```

```
Output \((\mu, p, t^*, \Pi)\)
```

38
D Proofs for Section 6

Proof of Proposition 3. We show the case where $J \subseteq J'$. The other case is symmetric. To simplify the notation, we denote the payoff for an agent $v \in I' \cup J'$ under $(\mu, t^*, \Pi)$ as $u(v)$ while the payoff under $(\mu', t'^*, \Pi')$ as $u'(v)$.

Suppose to the contrary that there exist $i \in I$ such that $u'(i) < u(i)$. Consider the sequence constructed by alternatively applying $\mu$ and $\mu'$ to $i$ until some agent is reached for a second time (shown in Figure 1), i.e., the sequence is either $i, \mu(i), \mu'(\mu(i)), \ldots, (\mu' \circ \mu)(i)$ or $i, \mu(i), \mu'(\mu(i)), \ldots, \mu \circ (\mu' \circ \mu)(i)$ for some integer $k$, which stops when encountering an existing agent. Obviously, the sequence has finitely many agents because $I' \cup J'$ is finite, which implies that $k$ is finite.

\[
\begin{array}{ccc}
  i & \mu(i) & \mu'(\mu(i)) \\
  \mu & \mu' & \mu'(\mu(i)) \\
  \mu & \mu & \mu \\
  \cdots & & \\
\end{array}
\]

Figure 1: An sequence by alternatively applying $\mu$ and $\mu'$ to $i$.

We claim that $u'(\mu(i)) > u(\mu(i))$. Otherwise, $(i, \mu(i))$ is a blocking pair for $(\mu', t', \Pi')$ because they have complete information about each other’s true type. Next we claim that $u'(\mu'(\mu(i))) < u(\mu'(\mu(i)))$. Otherwise, $(\mu'(\mu(i)), \mu(i))$ is a blocking pair for $\mu$, i.e., $\mu$ is not complete-information stable. We continue this argument until the end of the sequence. Particularly, either $u(\mu((\mu' \circ \mu)^{k-1}(i))) < u'((\mu' \circ \mu)^{k-1}(i)))$ or $u((\mu' \circ \mu)^{k}(i)) > u'((\mu' \circ \mu)^{k}(i)))$. Consider two cases.

Case 1. The sequence stops with $(\mu' \circ \mu)^{k}(i)$, i.e., firm $\mu((\mu' \circ \mu)^{k-1}(i)))$ is matched under $\mu'$ with some worker $(\mu' \circ \mu)^{k}(i)$ who was unmatched under $\mu$. Since $u(\mu((\mu' \circ \mu)^{k-1}(i))) < u'((\mu' \circ \mu)^{k-1}(i)))$, we know that $((\mu' \circ \mu)^{k}(i), \mu((\mu' \circ \mu)^{k-1}(i)))$ is a blocking pair for $\mu$, i.e., $\mu$ is not complete-information stable, a contradiction.

Case 2. The sequence stops with $\mu \circ (\mu' \circ \mu)^{k}(i)$, i.e., firm $\mu \circ (\mu' \circ \mu)^{k}(i)$ is unmatched under $\mu'$. Note that worker $(\mu' \circ \mu)^{k}(i)$ and firm $\mu((\mu' \circ \mu)^{k-1}(i))$ have complete information about each other’s true type. Since $u((\mu' \circ \mu)^{k}(i)) > u'((\mu' \circ \mu)^{k}(i)))$, we know that $((\mu' \circ \mu)^{k}(i), \mu((\mu' \circ \mu)^{k}(i)))$ is a blocking pair for $(\mu', t', \Pi')$, i.e., $(\mu', t', \Pi')$ is not a stable $\Gamma$-state, a contradiction.

Suppose to the contrary that there exists $j \in J$ such that $u'(j) > u(j)$. We have argued that $u'(i) \geq u(i)$ for all $i \in I$ and particularly $i = \mu^{-1}(j)$. Then $(\mu^{-1}(j), j)$ is a blocking pair for state prime since they have complete information about each other’s type.

\[\square\]
E  Proofs for Section 7

To prove Proposition 4, we first argue that for iteration (18) $\Sigma^{[T]+1} = \Sigma^{[T]}$, i.e., the sequence $\{\Sigma^k\}_k$ converges in finite time.

**Lemma 4.** $\Sigma^{[T]+1} = \Sigma^{[T]}$.

**Proof.** Let the set of individually rational outcomes associated with $(\mu, p)$ be defined as

$$\Sigma^0_{\mu, p} := \left\{ (\mu', p', t) \in \Sigma^0 : (\mu', p') = (\mu, p) \right\}.$$  

Similar to the iteration of (18), we could conduct iteration on the set $\Sigma^0_{\mu, p}$ as follows.

$$\Sigma^k_{\mu, p} := \left\{ (\mu, p, t) \in \Sigma^k_{\mu, p} : (\mu, p, t) \text{ is } \Sigma^k_{\mu, p} \text{ - stable} \right\}.$$  

It follows from induction that for each $k \geq 1$, $(\mu, p, t)$ is $\Sigma^k$-blocked if and only if $(\mu, p, t)$ is $\Sigma^k_{\mu, p}$-blocked. Therefore, we have

$$\Sigma^\infty = \bigcup_{(\mu, p) \in \mathcal{A}} \Sigma^k_{\mu, p}.$$  

The above equation implies that conducting the iteration process allocation-by-allocation is equivalent to conducting the original iteration (18); the predicted outcomes are the same.

Let $K_{\mu, p}$ be the *smallest* integer such that $\Sigma^{K_{\mu, p}+1}_{\mu, p} = \Sigma^{K_{\mu, p}}_{\mu, p}$. $K_{\mu, p}$ must be finite because $T \subset W^I \times F^J$ is a finite set and $\Sigma^{K_{\mu, p}+1}_{\mu, p}$ is a strict subset of $\Sigma^{K_{\mu, p}}_{\mu, p}$ if $\Sigma^{K_{\mu, p}+1}_{\mu, p} \neq \Sigma^{K_{\mu, p}}_{\mu, p}$. Moreover, $K_{\mu, p} \leq |T|$ for all $(\mu, p) \in \mathcal{A}$. Therefore, $\Sigma^{[T]+1} = \Sigma^{[T]}$.

**Lemma 5.** Suppose $(\mu, p, t) \in \Sigma \subset \Sigma'$. If $(\mu, p, t)$ is $\Sigma'$-blocked, then $(\mu, p, t)$ is $\Sigma$-blocked.

**Proof.** Since $(\mu, p, t)$ is $\Sigma'$-blocked, by Definition 6, $(\mu, p, t, \Pi_{\Sigma', \mu, p})$ is blocked. Lemma 2 implies that $(\mu, p, t, \Pi_{\Sigma', \mu, p, \Sigma'_{\mu, p}})$ is blocked because $\Pi_{\Sigma', \mu, p, \Sigma'_{\mu, p}}$ is finer than $\Pi_{\Sigma', \mu, p}$. Note that $\Pi_{\Sigma', \mu, p}(t') = [\Pi_{\Sigma', \mu, p, \Sigma'_{\mu, p}}](t')$ for all $v \in I \cup J$ and $t' \in T_{\Sigma', \mu, p}(t)$. Therefore, $(\mu, p, t, \Pi_{\Sigma', \mu, p})$ is blocked if and only if $(\mu, p, t, \Pi_{\Sigma', \mu, p, \Sigma'_{\mu, p}})$ is blocked. Hence $(\mu, p, t)$ is $\Sigma$-blocked.

**Proof of Proposition 4.** 1. By definition, $E \subset \Sigma^0$. Suppose $E \subset \Sigma^{k-1}$, for $k \geq 1$, and $(\mu, p, t) \in E$. Since $E$ is self-stabilizing, $(\mu, p, t)$ is $E$-stable, and so is $\Sigma^{k-1}$-stable (by Lemma 5 because $\Sigma^{k-1}$ is a larger set), and so is $\Sigma^k$ by the definition of $\Sigma^k$. Induction shows that $E \subset \Sigma^\infty$.

2. $\Sigma^\infty$ is self-stabilizing because $\Sigma^\infty = \Sigma^{[T]+1} = \Sigma^{[T]}$ (Lemma 4).
Proof of Proposition 5. Necessity. Suppose \((\mu, p, t) \in \Sigma^\infty\). Consider the partition profile \(\Pi_{\Sigma^\infty,\mu,p}\) defined as in (16)-(17). Since \((\mu, p, t)\) is not \(\Sigma^\infty\)-blocked, we know that \((\mu, p, t, \Pi_{\Sigma^\infty,\mu,p})\) is not blocked. Similarly, \((\mu, p, t', \Pi_{\Sigma^\infty,\mu,p})\) is not blocked for all \(t' \in T_{\Sigma^\infty,\mu,p}(t)\).

Notice that \(N(\mu,p,\Pi)(t) \supset T_{\Sigma^\infty,\mu,p}(t)\) by the definition of \(T_{\Sigma^\infty,\mu,p}(t)\). Then
\[
[H_{\mu,p}(\Pi_{\Sigma^\infty,\mu,p})]_k(t') = \Pi_k(t')
\]
for all \(t' \in T_{\Sigma^\infty,\mu,p}(t)\) and all \(v \in I \cup J\). Since \((\mu, p, t', \Pi_{\Sigma^\infty,\mu,p})\) is not blocked for all \(t' \in T_{\Sigma^\infty,\mu,p}(t)\), we know that \((\mu, p, t', H_{\mu,p}(\Pi_{\Sigma^\infty,\mu,p}))\) is not blocked for all \(t' \in T_{\Sigma^\infty,\mu,p}(t)\), which implies that \(N(\mu,p,H_{\mu,p}(\Pi_{\Sigma^\infty,\mu,p}))(t) \supset T_{\Sigma^\infty,\mu,p}(t)\). Keep applying this argument until we find a fixed point of \(H_{\mu,p}\). This must be done within finitely many times because the partition profile gets strictly finer whenever it is not a fixed point. Denote the fixed point by \(\Pi^*\). Then induction shows that \(N(\mu,p,\Pi^*)(t) \supset T_{\Sigma^\infty,\mu,p}(t)\) and that \((\mu, p, t', \Pi^*)\) is not blocked for all \(t' \in T_{\Sigma^\infty,\mu,p}(t)\), particularly for \(t' = t\).

Sufficiency. Suppose there exists a partition profile \(\Pi\) such that \((\mu, p, t, \Pi)\) is stable. Let \(\Sigma^N := \{(\mu, p, t') : t' \in N(\mu,p,\Pi)(t)\}\). Then obviously \((\mu, p, t) \in \Sigma^N\). By definition of \(N(\mu,p,\Pi), (\mu, p, t')\) is not \(\Sigma^N\)-blocked for all \(t' \in N(\mu,p,\Pi)(t)\). (Otherwise, \((\mu, p, t', \Pi)\) is blocked by Lemma 2, contradicting \(t' \in N(\mu,p,\Pi)(t)\).) Therefore, \(\Sigma^N\) is a self-stabilizing set. By Proposition 4, \((\mu, p, t) \in \Sigma^N \subset \Sigma^\infty\). \(\square\)

Proof of Corollary 2. Since \(T \subset T'\), we have \(\Sigma^0(T) \subset \Sigma^0(T')\). By Lemma 2, \(\Sigma^1(T) \subset \Sigma^1(T')\). Induction shows that \(\Sigma^\infty(T) \subset \Sigma^\infty(T')\). By Corollary 1, \(\mathcal{S}(T) \subset \mathcal{S}(T')\). \(\square\)