# Assignment Game with Private Payments

Mohsen Pourpouneh* Rasoul Ramezanian† Arunava Sen ‡

Vilok Taori§

September 15, 2023

**Abstract**

We adapt the Assignment Game model in Shapley and Shubik (1971) to study a stability notion in labour markets with private agreement regarding division of surplus between each matched pair. We propose a stability notion based on iterative elimination of blocked matching outcomes, which in addition to requiring individual rationality and no blocking pairs, captures the idea that absence of blocking pairs conveys no further information regarding the payoffs received by the other participants in the market. We define an algorithm to identify the set of stable matching outcomes (private payments stable set) and characterize it. We show that the set of stable matching outcomes exists, is efficient in terms of maximizing total surplus and includes the complete information stable matching outcomes.

## 1 Introduction

Consider decentralized labour markets where workers and firms match with each other. Each worker supplies a single unit of labour and each firm demands a single unit of labour. Each matched pair enters into a private agreement between them regarding division of surplus that is generated if they match together. What matching outcomes should we expect to persist in these markets?

*Postdoctoral Fellow, University of Oxford, Department of Computer Science

†Senior SNFS Research Manager, HEC, University of Lausanne

‡Professor, Economics and Planning Unit, Indian Statistical Institute, Delhi Centre

§Senior Research Fellow, Indian Statistical Institute, Delhi Centre

A large literature uses the matching models introduced by Gale and Shapley (1962) and Shapley and Shubik (1971) to analyze problems with two sided heterogeneity, studying problems such as matching men to women, and workers to firms. These models abstract away from the exact strategic interaction among participants in the market and use cooperative concept of stability that focuses on payoff assumptions. The framework assumes that agents fully observe an actual market outcome, in particular how the surplus is divided between each matched pair. Stability describes a situation in which there is no profitable coalition deviation from this outcome.

In this paper, we also abstract away from the exact strategic interaction of how the agents match and the exact bargaining protocol that results in a private agreement and propose a cooperative notion of stability. The paper examines matching models in which participants partially observe a matching outcome. In particular, both firms and workers can observe the assignment but they cannot observe the surplus division or the payment transfer between each matched pair apart from the pair where they are involved. We address the following questions. What does it mean for a matching outcome to be stable under incomplete information regarding payment transfer? What are the properties of the stable outcomes?

### 1.1 Beliefs

The key to our stability notion is specification of the beliefs of the agents regarding the payoff received by agents on the other side of the market to whom they may want to propose to block a candidate stable assignment. Our notion of stability is similar to that proposed by Liu et al. (2014) who study a matching environment where participants in the labour market fully observe actual market outcome but firms have incomplete information regarding worker types.

Consider a worker firm matching problem, in which each worker firm pair if matched together generate a surplus that is commonly known in the market. The participants in the market do not observe the division of surplus between the matched pair (i.e., the payment of wage from firm to worker). As in the complete information framework, we would say that the outcome is not stable if there is an unmatched worker firm pair that can deviate and increase payoff to each. But how does the firm (worker) estimate the worker's (firm's) payoff when proposing to block with the worker (firm) whose payoff is unknown in the given matching.

Suppose a firm is considering forming a blocking pair with the worker, we begin by identifying the payoffs of the worker that the firm can exclude, given the knowledge of the assignment and the hypothesis that the matching is not blocked. In particular, the firm may make inferences about the workers payoff from the lack of worker-firm pairs wishing to block. These inferences may lead to yet further inferences. This gives rise to firms' belief of the maximum payoff attained by the workers given that the matching is not yet blocked. We then say a matching outcome fails to be stable if some worker-firm pair has a deviation that is profitable, under firm's belief of the maximum payoff received by the worker given that the matching outcome is not yet blocked. Hence, we are considering the notion of blocking such that when either the firm or worker proposes according to their updated inference of the other party's payoff, it is certain that the other party will accept the proposal.

In motivating the final step we must distinguish between the viewpoint of either firm or worker and that of the analyst. The firm (worker) may have a particular belief drawn from the set of reasonable beliefs that depends on the actual strategic interaction among the participants in the market. However, nothing in the structure of the environment or the candidate stable allocation gives the analyst any clue as to what the firm's or the worker's belief might be. Our goal is to identify necessary conditions for stability that follow only from the structure of the environment and the hypothesis of stability and we accordingly reject an allocation only if we are certain that there is a successful block.

### 1.2 Preview

We discuss the related literature in Section 2. Sections 3 and 4 develop our stability concept for matching problems with private payments. We call this set of matching outcomes as the private payments stable set and prove that for any environment, this set of matching outcomes is non-empty.

Section 5 explores the implications of our stability notion. Efficiency of assignment in the sense of maximizing total surplus is a remarkable structural property of complete information stable matching outcomes. The set of matching outcomes in the private payments stable set are also efficient and are a super set of complete information stable outcomes. For the complete information stable set and private payments stable set, the maximum and minimum wage that each worker $i$ can attain are equal in both the sets. Hence, even in labour markets with private agreements, given our stability notion only efficient matching outcomes are expected to persist.

Section 6, using a numerical example we study the structure of private payment stable set and relate it to the complete information stable set (core).

## 2 Related Literature

The literature related to the core in assignment games started with the seminal work by Shapley and Shubik (1971) who established existence and proved that only the efficient assignments belong to the core. They also showed that the workers' imputations in every stable matching outcome correspond to their personalized walrasian prices. This literature assumed that there is complete information regarding agents payoff from a given matching outcome.

There has been a sizeable literature related to the use of theory of cooperative games to analyze situations with incomplete information. This literature has mainly analyzed adverse selection problems that is where the incomplete information is related to players' types. There has been two distinct strands in this literature.

The first is related to the study of incomplete information core that analyzes situations where the final outcome is not observed. This literature started with the path breaking work by Wilson (1978) who proposed "coarse core" and "fine core" corresponding to two polar protocols of information aggregation within a coalition. Dutta and Vohra (2005) proposed credible core in which coalitions are allowed to coordinate their objection by inferring from the objection being contemplated.

The other strand is related to matching models with incomplete information related to players' types. This literature analyzes a situation in which the final outcomes are fully observed and players consider deviating from this outcome based on their updated information and inference. This literature started with the seminal work by Liu et al. (2014) who considered matching environments where firms have incomplete information regarding worker types though there is complete information regarding payment transfer. They propose a stability notion with the requirement that firms make full use of the information they can infer from the common knowledge that the matching outcome is not blocked. Liu et al. (2014) show that under certain monotonicity and supermodualirity assumption regarding players value functions, every stable matching outcome is efficient. In our paper we consider matching environments where agents partially observe final matching outcome and the only incomplete information is related to this partial observation, that is the payment transfer. Our stability notion is similar in spirit to that of Liu et al. (2014). Chen and Hu (2020) extend the model of Liu et al. (2014) and prove that a random matching process leads to stable matching outcomes with probability one. Chen and Hu (2023) generalize the stability notion of Liu et al. (2014) by considering matching environments with two sided incomplete information and arbitrary information structures. Pomatto (2022) considers a non cooperative matching model and applies forward induction reasoning to arrive at the set of stable matching outcomes as in Liu et al. (2014). Bikhchandani (2017) considers a bayesian setting with nontransferable utilities and proposes a stability notion similar in spirit to Liu et al. (2014). Liu (2020) proposes a stability criterion that requires the Bayesian consistency of three beliefs; namely, the exogenously given prior beliefs, the off-path beliefs conditional on counterfactual pairwise blockings, and the on-path beliefs for stable matchings in the absence of such blockings.

## 3 Matching with Private Payments

### 3.1 Environment

There is a finite set of workers, indexed by $I$, with each individual worker denoted by $i \in I$. There is also a finite set of firms, indexed by $J$, with each individual firm denoted by $j \in J$. Without loss of generality we assume that $|I| \leq |J|$. A match between a worker $i$ and firm $j$ generates a surplus of $\nu_{ij}$.

An assignment between workers and firms is a bijection $\mu : I \longrightarrow J$. A matching outcome comprises of an assignment and a wage vector $(\mu, p)$, where the $i^{th}$ component of the wage vector $(p_{i\mu(i)})$ represents the transfer from firm $\mu(i)$ to worker $i$. Let $V$ be an $n \times m$ matrix, where $(i, j)^{th}$ component represents the surplus generated by worker $i$ and firm $j$. Each firm's and worker's index is commonly known, hence the surplus $(\nu_{i\mu(i)})$ generated by each matched pair given the assignment $\mu$ is common knowledge. All unmatched agents in a given assignment receive a payoff of 0. Given a matching outcome $(\mu, p)$, the wage vector $p$ is not common knowledge. Each worker $i$ has information regarding the wage received $p_{i\mu(i)}$ and each firm $j$ has information regarding the wage paid by him $p_{\mu^{-1}(j)j}$ given the assignment $\mu$.

Given a matching outcome $(\mu, p)$, then worker $i's$ payoff:

$$w_i^{(\mu,p)} = p_{i\mu(i)}$$

while the firm $\mu(i)'s$ payoff is

$$f_j^{(\mu,p)} = \nu_{i\mu(i)} - p_{i\mu(i)}$$

**Example 1** *Let* $I = \{i_1, i_2, i_3\}$ *be the set of workers and* $J = \{j_1, j_2, j_3\}$ *be the the set of firm. Let* $V$ *be the following matrix with the* $(i, j)^{th}$ *entry representing the surplus generated* *if worker $i$ and firm $j$ are matched together. Furthermore, consider the matching $(\mu, p)$ with $\mu = \{(i_1, j_2), (i_2, j_3), (i_3, j_1)\}$ and the wage vector $p$ with $p_{i_1 j_2} = 3$, $p_{i_2 j_3} = 2$, and $p_{i_3 j_1} = 0.05$.*

$$\begin{array}{cc} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ V = & \begin{bmatrix} 4 & 4.1 & 3.2 \\ 2.8 & 3 & 3.1 \\ 1.7 & 1.8 & 2 \end{bmatrix} \begin{array}{c} i_1 \\ i_2 \\ i_3 \end{array} \end{array}$$

The notion of stability for this complete information setting is familiar from Gale and Shapley (1962) and Shapley and Shubik (1971). For the above matching, worker $i_3$ knows that firm $j_3$ earns a payoff of 1.1. By forming a matching block with firm $j_3$, they can generate a surplus of 2 and make a payment under which both receive more than the candidate match. Therefore the matching $(\mu, p)$ in Example 1 is not complete information stable.

Now suppose that in the given matching, workers and firms know their own payoffs but do not know the payoffs of the other participants in the market, though they know the surplus generated by each matched pair. Consider a blocking proposal by worker $i_3$ of a candidate blocking pair with firm $j_3$ and some payment, $\tilde{p} \in (0.05, 0.1]$ . The proposed block is profitable for worker $i_3$. Worker $i_3$ does not know the payoff earned by firm $j_3$ in its matching with worker $i_2$. As worker $i_3$ knows that firm $j_3$ is matched with firm $i_2$, worker $i_3$ can infer that the maximum payoff earned by firm $j_3$ cannot exceed 3.1. With this information, $i_3$ cannot be certain that firm $j_3$ will accept his proposal. As the information that the matching outcome is not blocked is not used to draw further inferences regarding firm $j_3$'s payoff, this is not a "reasonable belief". Worker $i_3$'s beliefs should be consistent with all inferences he can draw using information of the environment and the hypothesis that the matching is not blocked. Given the hypothesis that the matching is not blocked, as firm $j_1$ can earn a maximum payoff of 1.7, worker $i_3$ infers that worker $i_1$ should receive a payoff of at least 2.3 otherwise he can form a blocking pair with firm $j_1$. This means that firm $j_2$ can earn a maximum payoff of 1.8. Worker $i_3$ infers that worker $i_2$ infers that firm $j_2$ can earn a maximum payoff of 1.8. Hence worker $i_3$ infers that worker $i_2$ should earn a payoff of at least 1.2, otherwise worker $i_2$ can form a blocking pair with firm $j_2$. Hence, worker $i_3$ infers that firm $j_3$ can earn a maximum payoff of 1.9. With this "reasonable belief" worker $i_3$ is certain that the blocking proposal will be accepted by firm $j_3$. Hence the matching $(\mu, p)$ is not also not private payments stable

# 4 Stability

## 4.1 Individual Rationality

An outcome $(\mu, p)$ is individually rational if each agent receives a payoff of at least zero. As the workers and firms know their own payoffs in a given matching, the notion of individual rationality is the same for complete and incomplete information.

**Definition 1** *A matching outcome $(\mu, p)$ is individually rational if for all workers $i \in I$ and all firms $j \in J$, we have:*

$$w_i^{(\mu,p)} = p_{i\mu(i)} \geq 0 \qquad \text{and} \qquad f_j^{(\mu,p)} = \nu_{\mu^{-1}(j)j} - p_{\mu^{-1}(j)j} \geq 0$$

## 4.2 Complete Information Stability

The notion of stability in matching games with transferable utilities was first formulated by Becker (1973) and Shapley and Shubik (1971) who also established existence and showed that the stable allocation is efficient in the sense of maximizing the total surplus. In what follows we summarize the notation and results regarding the matchings in complete information setting.

**Definition 2** *A matching outcome $(\mu, p)$ is complete information stable, if it is individually rational and there is no unmatched worker-firm combination $(i, j)$ and a payment $p_{ij}$ such that:*

$$p_{ij} \geq p_{i\mu(i)} \qquad \textit{and} \qquad \nu_{ij} - p_{ij} \geq \nu_{\mu^{-1}(j)j} - p_{\mu^{-1}(j)j} \tag{1}$$

*with strictly higher payoff for either worker $i$ or firm $j$.*

Note that, adding the corresponding sides in the inequality 1, implies that a matching outcome $(\mu, p)$ is complete information stable, whenever there is no unmatched worker-firm combination $(i, j)$ such that:

$$\nu_{ij} > w_i^{(\mu,p)} + f_j^{(\mu,p)} \tag{2}$$

**Definition 3** *Given surplus matrix $V$, an assignment $\mu$ is efficient whenever for every other assignment $\mu' \neq \mu$,*

$$\sum_{i=1}^{n} \nu_{i\mu(i)} \geq \sum_{i=1}^{n} \nu_{i\mu'(i)} \tag{3}$$

In words, an assignment is efficient whenever it maximizes the total surplus. Throughout this paper we denote the set of efficient assignments for a given surplus matrix $V$ by $Ef(V)$.

**Result 1** *(Stable matching Shapley and Shubik (1971)). A matching is complete information stable if and only if it is efficient. Equivalently, an assignment is complete information stable if and only if it is the solution to the linear programming problem* $\max_{\mu \in \mathcal{M}} \sum_{i \in I} \nu_{i\mu(i)}$*, where* $\mathcal{M}$ *is the set of feasible matching.*

**Result 2** *(Full matching Shapley and Shubik (1971)). If every element in the surplus matrix is positive, a complete information stable matching is a full matching; that is, the number of matched pairs in the complete information stable outcome reaches the maximal possible number.*

**Result 3** *(Efficient matching Shapley and Shubik (1971)). If there is a unique efficient matching, this matching is the unique matching in the complete information stable outcome.*

**Result 4** *(Lattice structure Shapley and Shubik (1971)). The set of complete information stable matching outcomes form a complete lattice. In particular, there is worker optimal stable matching outcome and there is a firm optimal stable matching outcome.*

### 4.3 Incomplete Information

We are interested in the stability of a matching when each agent knows his own payoff but not the payoffs of the other agents, though each worker and firm can observe the surplus generated by each matched pair. We view stability as capturing a notion of steady state: a matching is stable if once established, it remains in place. Think of workers and firms in the labor market observing a particular matching. If the matching is stable, then we should expect to see the the same matching when next time the labour market opens and each subsequent time the labor market opens. We model firms' and workers' inferences of the other agents' payoff using iterated elimination of blocked matching outcomes. This procedure is related to game theoretic notion of rationalizability by Bernheim (1984) and Pearce (1984) of eliminating strategies that are never best responses.

Consider a firm contemplating a blocking match with a worker knowing that realization of worker's payoffs is consistent with set of matching outcomes $\Sigma$. Our notion of blocking is designed to only exclude outcomes that we are certain will give rise to objection.

**Definition 4** *Fix a non empty subset of individually rational matching outcomes* $\Sigma$*. We say a matching outcome* $(\mu, p) \in \Sigma$ *is* $\Sigma$*-blocked if there is a worker-firm pair* $(i, j)$ *and payment* $p_{ij}$

*such that either for all* $(\mu, \bar{p}) \in \Sigma$,

$$p_{ij} \geq p_{i,\mu(i)} \qquad \textit{and} \qquad \nu_{ij} - p_{ij} \geq \nu_{\mu^{-1}(j)j} - \bar{p}_{\mu^{-1}(j)j} \tag{4}$$

*or for all* $(\mu, \bar{p}) \in \Sigma$,

$$p_{ij} \geq \bar{p}_{i,\mu(i)} \qquad \textit{and} \qquad \nu_{ij} - p_{ij} \geq \nu_{\mu^{-1}(j)j} - p_{\mu^{-1}(j)j} \tag{5}$$

where in each inequality 4 and 5 at least one of the inequalities must be strict. The inequality 4 in the above definition relates to the case when worker $i$ proposes to firm $j$. The former inequality (in 4) requires that worker $i$ receives a higher payoff in the proposed block with firm $j$ than in the current matching. The latter inequality (in 4) requires that under any reasonable belief that worker $i$ has regarding firm $j's$ payoff, firm $j$ will accept the proposed block with worker $i$. By reasonable beliefs, we mean worker $i's$ belief regarding firm $j's$ payoffs that must be consistent with matching outcomes in the set $\Sigma$, a restriction that will become operational in the iterative argument we construct next.

Let $\hat{f}_j^{(\mu,\Sigma)}$ and $\hat{w}_i^{(\mu,\Sigma)}$ be the maximum payoff firm $j$ and worker $i$ receive in assignment $\mu$ given the set of matching outcomes $\Sigma$, respectively. Adding the inequalities in 4 and the inequalities in 5, we have an alternative definition of a matching outcome $(\mu, p)$ being $\Sigma$-blocked.

**Definition 5** *A matching outcome* $(\mu, p)$ *is* $\Sigma$*-blocked if there exists worker* $i$ *and firm* $j$ *such that either*

$$\nu_{ij} > w_i^{(\mu,p)} + \hat{f}_j^{(\mu,\Sigma)} \tag{6}$$

*or*

$$\nu_{ij} > \hat{w}_i^{(\mu,\Sigma)} + f_j^{(\mu,p)} \tag{7}$$

*We say an individually rational matching outcome is* $\Sigma$*-stable if it is not* $\Sigma$*-blocked.*

**Remark.** For a given matching that is $\Sigma$-stable, neither the worker nor the firm has incentive to propose to other firms or worker. We refer to the case that the workers have incentive to propose as the $\Sigma(w)$-blocked, i.e., the case corresponding to Inequality 6. Similarly, we define $\Sigma(f)$-blocked to refer to the case that firms have incentive to propose, i.e., the case corresponding to Inequality 7.

How does an allocation and its immunity to blocking be commonly known. Our view is that stable matching is one that we should expect to persist and we thus think of agents in the market repeatedly observing this outcome. We assume that all agents are rational and there is common belief in rationality, hence each time they partially observe the matching outcome, they can draw inferences about its properties- first that is individually rational, then that everyone knows it is individually rational and there are no blocking pairs, then everyone knows that and so on. Each observation corresponds to a step in the iterative belief process until no further information can be inferred from individual rationality and absence of blocking.

**Definition 6** *Let $\Sigma^0$ be set of all individually rational matching outcomes. For $k \geq 1$, let*

$$\Sigma^k = \Big\{ (\mu, p) \in \Sigma^{k-1} : (\mu, p) \quad is \quad \Sigma^{k-1} - stable \Big\}.$$

*The set of private payments stable matching outcomes is given by $\Sigma^* = \bigcap_{k=0}^{\infty} \Sigma^k$.*

The set $\Sigma^k$ is a (weakly) decreasing sequence of set of matching outcomes. The next proposition shows that the set of private payments stable matchings $\Sigma^*$ is non-empty.

**Proposition 1** *For every surplus matrix $V$, the set of private payment stable matching outcomes $\Sigma^*$ is non empty.*

**Proof** : In Shapley and Shubik (1971), it is shown that for any surplus matrix $V$, there exist a complete information stable matching $(\mu, p)$, which is efficient and maximizes the total surplus. As $(\mu, p)$ is complete information stable then for every worker $i$ and firm $j$, Definition 2 implies,

$$w_i^{(\mu,p)} + f_j^{(\mu,p)} \geq \nu_{ij}$$

Since, $\hat{f}_j^{(\mu,\Sigma)}$ and $\hat{w}_i^{(\mu,\Sigma)}$ are the maximum payoffs of firm $j$ and worker $i$ in assignment $\mu$ with respect to the set of matching outcomes $\Sigma$, then for every $k \geq 0$, we have $\hat{f}_j^{(\mu,\Sigma^k)} \geq f_j^{(\mu,p)}$ and $\hat{w}_i^{(\mu,\Sigma^k)} \geq w_i^{(\mu,p)}$. This implies that for every worker $i$ and firm $j$ and for every $k \geq 0$

$$w_i^{(\mu,p)} + \hat{f}_j^{(\mu,\Sigma^k)} \geq w_i^{(\mu,p)} + f_j^{(\mu,p)} \geq \nu_{ij}$$

and

$$\hat{w}_i^{(\mu,\Sigma^k)} + f_j^{(\mu,p)} \geq w_i^{(\mu,p)} + f_j^{(\mu,p)} \geq \nu_{ij}$$

Therefore by Definition 5, $(\mu, p) \in \Sigma^*$. ■

## 5 Private Payment Stable Set

Our aim is to identify the private payment stable set of matching outcomes for any arbitrary surplus matrix $V$. For each assignment $\mu$, we determine the set of wage vectors $p$ such that $(\mu, p) \in \Sigma^*$. Thought the rest of this paper, without loss of generality, we assume the workers and firms are relabeled such that $\mu(i) = i$.

The following procedure determines the set of all wage vectors that are prevents any worker from proposing a blocking pair with any other firm. For the given assignment $\mu$, the maximum payoff that can be received by each firm $i$, among the set of all individually rational matching outcomes $\Sigma^0$, is common knowledge among the workers and it is equal to $\hat{f}_i^{(\mu,\Sigma^0(w))} = \nu_{ii}$ for all $i \in N$. Let $A^1$ be an $n \times n$ matrix where the $(i,j)^{th}$ entry represents the maximum individually rational payoff that worker $i$ receives if he forms a blocking pair with firm $j \neq i$, provided firm $j$ receives a payoff greater than or equal to $\hat{f}_i^{(\mu,\Sigma^0(w))}$. Formally,

$$a_{ij}^1 = \begin{cases} 0, & \text{if} \quad i = j, \\ \max(\nu_{ij} - \hat{f}_j^{(\mu,\Sigma^0(w))}, 0) & \text{if} \quad i \neq j. \end{cases}$$

Let $a^1$ be a column matrix that denotes the maximum of each row in the $A^1$ matrix. Formally, with slight abuse of notation, let $a_i^1 = \max_{j \neq i} a_{ij}^1$. Proceeding recursively, for iteration $k \geq 1$, it is common knowledge among workers that the set of matching outcomes not yet blocked is $\Sigma^k(w)$. The maximum payoff that can be received by each firm $i$ in assignment $\mu$, given the set of unblocked matching outcomes $\Sigma^k(w)$ is updated:

$$\hat{f}_i^{(\mu,\Sigma^k(w))} = \nu_{ii} - a_i^k$$

Similarly, let $A^{k+1}$ be an $n \times n$ matrix where the $(i,j)^{th}$ entry $a_{ij}^{k+1}$ represents the maximum individually rational payoff that worker $i$ receives if he forms a blocking pair with firm $j$ ($j \neq i$), provided firm $j$ receives a payoff greater than or equal to $\hat{f}_i^{(\mu,\Sigma^k(w))}$. Formally, $a_{ij}^{k+1}$ is defined as:

$$a_{ij}^{k+1} = \begin{cases} 0, & \text{if} \quad i = j, \\ \max(\nu_{ij} - \hat{f}_j^{(\mu,\Sigma^k(w))}, 0) & \text{if} \quad i \neq j. \end{cases}$$

The above procedure eliminates wage vectors that cannot support assignment $\mu$ from the perspective of workers to form a blocking pair with firms based on their inference from observing the matching outcomes that are not yet blocked. To see this, consider any $k \geq 0$, and let $(\mu, p)$ be a matching outcome for which there exist a worker $i$ and firm $j$ with $p_{ii} < \max\limits_{j \neq i} a_{ij}^{k+1}$. As

$w_i^{(\mu,p)} = p_{ii}$, this implies that $\nu_{ij} > p_{ii} + \hat{f}_j^{(\mu,\Sigma^k(w))} = w_i^{(\mu,p)} + \hat{f}_j^{(\mu,\Sigma^k(w))}$, which (according to Definition 5 is a $\Sigma^k(w)$-blocked). Therefore, for such wage vectors $(\mu, p) \notin \Sigma^{k+1}(w)$.

Similarly, we can define the firms proposing procedure and and the eliminating the wage vectors from the perspective of firms. That is, for the given assignment $\mu$, the maximum payoff that can be received by each worker $i$, among the set of all individually rational matching outcomes $\Sigma^0(f)$, is common knowledge among the firms and it is equal to $\hat{w}_i^{(\mu,\Sigma^0(f))}$ for all $i \in N$. Formally,

$$b_{ij}^1 = \begin{cases} 0, & \text{if} \quad i = j, \\ \max(\nu_{ij} - \hat{w}_i^{(\mu,\Sigma^0(f))}, 0) & \text{if} \quad i \neq j. \end{cases}$$

Let $b^1$ be a row matrix that denotes the maximum of each column in the $B^1$ matrix. Formally, with slight abuse of notation, let $b_j^1 = \max_{j \neq i} b_{ji}^1$. Proceeding recursively, for iteration $k \geq 1$, it is common knowledge among firms that the set of matching outcomes not yet blocked is $\Sigma^k(f)$. The maximum payoff that can be received by each worker $i$ in assignment $\mu$, given the set of unblocked matching outcomes $\Sigma^k(f)$ is updated:

$$\hat{w}_i^{(\mu,\Sigma^k(f))} = \nu_{ii} - b_i^k$$

Similarly, let $B^{k+1}$ be an $n \times n$ matrix where the $(i,j)^{th}$ entry $b_{ij}^{k+1}$ represents the maximum individually rational payoff that firm $j$ receives if he forms a blocking pair with worker $i$, provided worker $i$ receives a payoff greater than or equal to $\hat{w}_i^{(\mu,\Sigma^k(f))}$. Formally, $b_{ij}^{k+1}$ is defined as:

$$b_{ij}^{k+1} = \begin{cases} 0, & \text{if} \quad i = j, \\ \max(\nu_{ij} - \hat{w}_i^{(\mu,\Sigma^k(f))}, 0) & \text{if} \quad i \neq j. \end{cases}$$

Similar to the argument for the workers, the above procedure eliminates wage vectors that cannot support assignment $\mu$ from the perspective of firms to form a blocking pair with workers based on their inference from observing the matching outcomes that are not yet blocked.

The set of $\Sigma$-stable, matchings are the set of those that are neither $\Sigma(w)$-blocked, nor $\Sigma(f)$-blocked. Therefore, the set of $\Sigma$-stable matchings are the ones at the intersection of the two.

**Example 2** *Consider the surplus matrix of Example 1. Consider the the matching $\mu(i) = i$.*

$$\begin{array}{cc} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ V = & \begin{bmatrix} 4 & 4.1 & 3.2 \\ 2.8 & 3 & 3.1 \\ 1.7 & 1.8 & 2 \end{bmatrix} \begin{array}{c} i_1 \\ i_2 \\ i_3 \end{array} \end{array}$$

*Given $\mu$, it is common knowledge among workers, the maximum each firm $i$ can receive given the set of all individually rational matching outcomes $\Sigma^0(w)$ is:*

$$\hat{f}^{(\mu,\Sigma^0(w))} = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$$

*Given $\hat{f}^{(\mu,\Sigma^0(w))}$, we can now determine minimum individually rational wage each worker $i$ should receive in $\mu$*

$$A^1 = \begin{bmatrix} 0 & 1.1 & 1.2 \\ 0 & 0 & 1.1 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow a^1 = \begin{bmatrix} 1.2 \\ 1.1 \\ 0 \end{bmatrix}$$

*Hence, for matching outcome $(\mu, p)$ to be not blocked, worker 1 must receive a wage of at least 1.2 (otherwise he can block with firm 3), and worker 2 must receives a wage of at least 1.1 (otherwise he can block with firm 3).*

*Given $a^1$, it is common knowledge among workers, the maximum each firm $i$ can receive given the set of matching outcomes not to be blocked is given by,*

$$\hat{f}^{(\mu,\Sigma^1(w))} = \begin{bmatrix} 2.8 \\ 1.9 \\ 2 \end{bmatrix}$$

*Given $\hat{f}^{(\mu,\Sigma^1(w))}$, we can now determine minimum wage each worker $i$ should receive in $\mu$.*

$$A^2 = \begin{bmatrix} 0 & 2.2 & 1.2 \\ 0 & 0 & 1.1 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow a^2 = \begin{bmatrix} 2.2 \\ 1.1 \\ 0 \end{bmatrix}$$

*Hence, for matching outcome $(\mu, p)$ to be not blocked, worker 1 must receive a wage of at least 2.2 (otherwise he can block with firm 2), and worker 2 must receive a wage of at least 1.1 (otherwise he can block with firm 3).*

*Given $a^2$, it is common knowledge among workers, the maximum each firm $i$ can receive given the set of matching outcomes not to be blocked is given by,*

$$\hat{f}^{(\mu,\Sigma^2(w))} = \begin{bmatrix} 1.8 \\ 1.9 \\ 2 \end{bmatrix}$$

*Given $\hat{f}^{(\mu,\Sigma^2(w))}$, we can now determine minimum wage each worker $i$ should receive in $\mu$,*

$$A^3 = \begin{bmatrix} 0 & 2.2 & 1.2 \\ 1 & 0 & 1.1 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow a^3 = \begin{bmatrix} 2.2 \\ 1.1 \\ 0 \end{bmatrix}$$

*As $a^2 = a^3$, the minimum wage each worker $i$ should receive does not change, hence the algorithm terminates.*

*Similarly, we consider firms proposing and forming a blocking pair with workers. Following the steps, we get the minimum payoff of the firms to be*

$$\begin{bmatrix} 0 & 0.1 & 0.2 \end{bmatrix}.$$

*This implies the set $\mathcal{P} = \{p : 2.2 \leq p_1 \leq 4,\ 1.1 \leq p_2 \leq 2.9,\ 0 \leq p_3 \leq 1.8\}$ can support the assignment $\mu$ as private payment stable outcome. Note that, the wage vector $\bar{p}$ with $\bar{p}_{i_1 j_1} = 4$, $\bar{p}_{i_2 j_2} = 2$, and $\bar{p}_{i_3 j_3} = 1$, with the same assignment $\mu$ i.e., $(\mu, \bar{p})$, is not complete information stable (since $i_3$ and $j_1$ can form a blocking pair), however it is private payments stable according to the above reasoning.*

Given assignment $\mu$, for any iteration $k \geq 1$ that is associated with the set of unblocked matching outcomes, let $a_i^k = \max_{j \neq i} a_{ij}^k$. This represent the minimum payoff each worker $i$ should receive given iteration $k$, whereas $\hat{w}_i^{(\mu,\Sigma^k(f))}$ represents the maximum payoff each worker $i$ can receive given iteration $k$.

**Proposition 2** *Let $V$ be an $n \times n$ surplus matrix. If $\mu \in Ef(V)$, then the above procedure terminates in at most $n$ iteration.*

**Proof :** See Appendix 7.1. ■

Shapley and Shubik (1971) showed that set of complete information stable matching outcomes form a complete lattice. In particular, there is worker optimal stable matching outcome and there is a firm optimal stable matching outcome. Let $\overline{w_i}$ and $\overline{f_i}$ be worker $i's$ and firm $i's$ payoff respectively in the worker optimal stable matching and $\underline{w_i}$ and $\underline{f_i}$ be worker $i's$ and firm $i's$ payoff respectively in the firm optimal stable matching.

Next we show that the output of the iterative elimination procedure for the workers corresponds to the firm optimal stable matching.

**Proposition 3** *Let $V$ be an $n \times n$ surplus matrix. If $\mu \in Ef(V)$, then the output of the iterative elimination procedure by workers is $\underline{w_i}$ for all workers $i$.*

**Proof :** By Proposition (2) the iterative elimination procedure by workers will terminate in $n$ steps. Let $a_i^n$ denote the wage of worker $i$ as the output of the procedure, and $a^n$ be the wage vector of the workers. We first show that $(\mu, a^n)$ is complete information stable. Note that, as $(\mu, a^n) \in \Sigma^*$, then for every worker $i$ and firm $j$, Definition 5 Inequality 6 implies $a_i^n + \hat{f}_j^{(\mu,\Sigma^k(w))} \geq \nu_{ij}$, which yields $a_i^n + (\nu_{jj} - a_j^n) \geq \nu_{ij}$. Then Definition 2 implies $(\mu, a^n)$ is complete information stable. As $\underline{w_i}$ is worker $i$'s minimum payoff among the set of complete information stable matching outcomes, hence for all workers $i$,

$$a_i^n \geq \underline{w_i}$$

By Shapley and Shubik (1971), the matching $(\mu, \underline{w_i})$ is complete information stable. Therefore, Proposition 1 implies $(\mu, \underline{w_i})$ is also private payment stable, i.e., $(\mu, \underline{w_i}) \in \Sigma^*$. By construction of the iterative elimination procedure by workers for any $(\mu, p) \in \Sigma^*$, we have $a_i^n \leq p_i$ for all workers $i$. In particular, as $(\mu, \underline{w_i}) \in \Sigma^*$, then

$$a_i^n \leq \underline{w_i}.$$

The last two inequalities implies $a_i^n = \underline{w_i}$ for all workers $i$. ■

**Proposition 4** *Let $V$ be an $n \times n$ surplus matrix. If $\mu \in Ef(V)$, then the output of the iterative elimination procedure by firms is $\overline{w_i}$ for all workers $i$.*

**Proof :** The proof follows from a similar structure to that of Proposition 3.

■

Putting Proposition 3 and 4 together we have,

**Corollary 1** *Let $V$ be an $n \times n$ surplus matrix. If $\mu \in Ef(V)$, then $(\mu, p) \in \Sigma^*$ if for every worker $i$,*

$$\underline{w_i} \leq p_i \leq \overline{w_i}$$

In Proposition 5, we show that for any assignment $\mu \notin Ef(V)$, there is no wage vector that can support it as a private payment stable matching outcome. To do so, we consider a restricted version of the iterative elimination procedure to eliminate wage vectors that cannot support the assignment $\mu$. In particular, we only allow workers to consider potential blocking pairs with firms whom they are assigned in a (particular) efficient assignment. We show that given this restricted procedure, there is always an agent that can violate the definition of private payment stability.

Let $\mu \notin Ef(V)$ and $\tilde{\mu} \in Ef(V)$ be an efficient assignment. Without loss of generality, we can relabel the index of workers and firms such that $\tilde{\mu}(i) = i$, then we have

$$\sum_{i=1}^{n} \nu_{ii} - \sum_{i=1}^{n} \nu_{\mu^{-1}(i)i} = \epsilon > 0 \tag{8}$$

For every worker $i$, let $f_i^0 = \nu_{\mu^{-1}(i)i}$, denote the maximum payoff received by each firm $i$ given assignment $\mu$ and given the set of all individually rational matching outcomes $\Sigma^0$. Note that $f_i^0$ is common knowledge among the workers in $N \setminus \{i\}$.

We define $\tilde{a}_i^1$ as the maximum payoff that worker $i$ receives if he forms a blocking pair with firm $i$ (partner in the efficient assignment), provided firm $i$ receives a payoff greater than $f_i^0$. That is, $\tilde{a}_i^1 = \max(\nu_{ii} - f_i^0, 0) = \max(\nu_{ii} - \nu_{\mu^{-1}(i)i}, 0)$. Given $\tilde{a}_i^1$, which is the lower bound for the payoff that worker $i$ should receive in assignment $\mu$ for it to be not blocked, the maximum each firm $i$ can receive in $\mu$ given this restricted blocking procedure is updated accordingly as, $f_i^1 = \nu_{\mu^{-1}(i)i} - \tilde{a}_{\mu^{-1}(i)}^1$.

Proceeding recursively, for iteration $k$, $\tilde{a}_i^k$ is the lower bound for the payoff that worker $i$ should receive in assignment $\mu$:

$$\tilde{a}_i^k = \max(\nu_{ii} - f_i^{k-1}, 0)$$

Hence, the maximum each firm can receive in $\mu$ is updated accordingly as:

$$f_i^k = \nu_{\mu^{-1}(i)i} - \tilde{a}_{\mu^{-1}}^k$$

The process proceeds recursively using this restricted procedure to eliminate wage vectors. For assignment $\mu$, if the wage vector $p$ can be eliminated by this procedure, then it is also eliminated by the general procedure as now each worker considers forming a blocking pair only with the firm that he is assigned in the efficient assignment.

We will use the following lemma to prove the main proposition.

**Lemma 1** *Given assignment $\mu$, if for some worker $i$ the minimum payoff that worker $i$ should receive in iteration $k$, exceeds the maximum surplus available to the worker-firm pair $(i, \mu(i))$, then there exists no payment vector $p$ such that $(\mu, p)$ is private payment stable.*

**Proof** : Let for worker $i$ and iteration $k$, the minimum payoff of worker $i$, exceeds the maximum surplus available to the worker-firm pair $(i, \mu(i))$, that is $\tilde{a}_i^k = \nu_{ii} - f_i^{k-1} > \nu_{i\mu(i)}$. Note that, $f_i^{k-1}$ gives upper bound for the payoff of firm $i$ in assignment $\mu$, i.e., in the restricted iterative elimination procedure, therefore $\hat{f}_i^{(\mu,\Sigma^{k-1}(w))} \leq f_i^{k-1}$, and $a_i^k \geq \tilde{a}_i^k$. Therefore,

$$a_i^k \geq \tilde{a}_i^k = \nu_{ii} - f_i^{k-1}(\mu) > \nu_{i\mu(i)} \Rightarrow \nu_{ii} > \nu_{i\mu(i)} + f_i^{k-1} \geq \nu_{i\mu(i)} + \hat{f}_i^{(\mu,\Sigma^{k-1}(w))}$$

Therefore, $\nu_{ii} > \nu_{i\mu(i)} + \hat{f}_i^{(\mu,\Sigma^{k-1}(w))}$. For any payment vector $p$, the maximum that the worker $i$ can get is the total surplus. That is, $\nu_{i\mu(i)} \geq p_i = w_i^{(\mu,p)}$. Putting this together with the latter inequality, $\nu_{ii} > w_i^{(\mu,p)} + \hat{f}_i^{(\mu,\Sigma^{k-1}(w))}$. However, by Inequality 6, worker $i$ and firm $i$ can form a blocking pair. Therefore there exists no payment vector that can support $\mu$ as a private payment stable matching outcome. ■

Finally, given Lemma 1 and the restricted elimination procedure we show,

**Proposition 5** *Suppose $\mu \notin Ef(V)$, then there exists no wage vector $p$ such that $(\mu, p) \in \Sigma^*$.*

**Proof** : See Appendix 7.2. ■

Using Corollary 1 and Proposition 5, we can state our main theorem:

**Theorem 1** *For any arbitrary surplus matrix $V$, a matching outcome $(\mu, p)$ is private payments stable, i.e., $(\mu, p) \in \Sigma^*$ if and only if*

$$\mu \in Ef(V), \quad \textit{and}$$
$$\underline{w_i} \leq p_i \leq \overline{w_i} \quad \textit{for every worker } i.$$

## 6 Structure of Private Payments Stable Set

The complete information stable set of any assignment game is rarely singleton. Consequently, the wage vector associated with efficient assignment in stable matching outcome is not unique. The set of wage vectors associated with complete information stable set is a convex polyhedron and is a subset of wage vectors associated with the private payments stable set. Importantly, the wage vector associated with worker optimal stable matching and firm optimal stable matching prove to be at least as far apart as any other two points in both these sets.

Consider the numerical example as in page 122 of Shapley and Shubik (1971) to understand the difference between the complete information stable set and the private payments stable set.

Suppose there are 3 workers and 3 firms. Let $V$ be the following $3 \times 3$ matrix with the $(i,j)^{th}$ entry representing the surplus generated if worker $i$ and firm $j$ are matched together.

$$V = \begin{array}{c} \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{bmatrix} 5 & 8 & 2 \\ 7 & 9 & 6 \\ 2 & 3 & 0 \end{bmatrix} \end{array} \begin{array}{c} \\ i_1 \\ i_2 \\ i_3 \end{array}$$

The unique efficient matching $\mu = \{(i_1, j_2), (i_2, j_3), (i_3, j_1)\}$. The wage vector associate with worker optimal stable matching is $w^* = (5, 6, 1)$ whereas the wage vector associated with firm optimal stable matching is $w_* = (3, 5, 0)$. The structure of the complete information stable set and private payments stable set is shown in Figure 1 and 2, respectively. The private payments stable set is a super set of the complete information stable set and the two wage vectors $w^*$ and $w_*$ are at least as far apart as any other two points in both these sets.

For any matching outcome in the private payments stable set that is not in the complete information stable set, all agents know that the assignment is associated with the unique efficient assignment but they do not know the payment transfer apart from the match where they are involved. Each worker knows that the maximum each firm can earn is equal to its payoff in the firm optimal stable matching and each firm knows that the maximum each worker can earn is equal to his payoff in the worker optimal stable matching. With this information, it is not certain for an analyst to infer that the matching outcome will be blocked by some worker firm pair. If the wage vectors were known to all, then it is certain that this matching outcome is blocked by a worker firm pair. The private payments stable set is a self stabilizing set and hence is a super set of the complete information stable set.

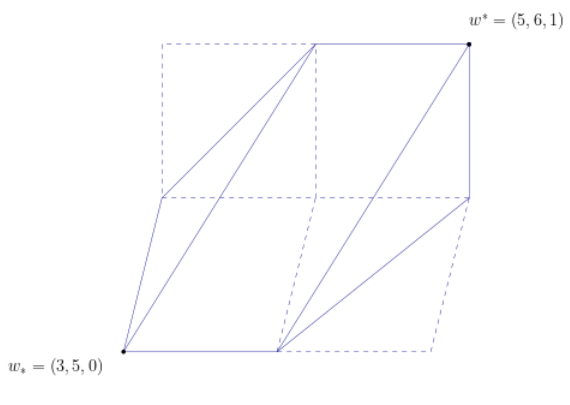


Figure 1: Complete information stable set.

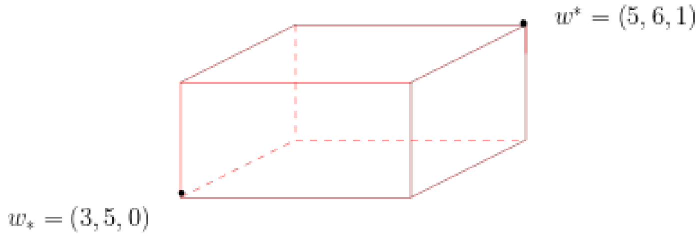


Figure 2: Private payments stable set.

## 7 Conclusion

In this paper, we propose a stability notion for matching in labour markets with private agreements regarding division of surplus. We show that the assignments corresponding to the matching outcomes in the private payments stable set are efficient like the case of complete information. As the wage vectors associated with the worker optimal and firm optimal stable matching correspond to their personalized walrasian prices, our stability notion provides a foundation to the existence of these in the labour markets.

## References


Becker, G. S. (1973). A theory of marriage: Part I. *Journal of Political economy*, 81(4):813–846.

Bernheim, B. D. (1984). Rationalizable strategic behavior. *Econometrica: Journal of the Econometric Society*, pages 1007–1028.

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

Chen, Y.-C. and Hu, G. (2020). Learning by matching. *Theoretical Economics*, 15(1):29–56.

Chen, Y.-C. and Hu, G. (2023). A theory of stability in matching with incomplete information. *American Economic Journal: Microeconomics*, 15(1):288–322.

Dutta, B. and Vohra, R. (2005). Incomplete information, credibility and the core. *Mathematical Social Sciences*, 50(2):148–165.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15.

Liu, Q. (2020). Stability and bayesian consistency in two-sided markets. *American Economic Review*, 110(8):2625–2666.

Liu, Q., Mailath, G. J., Postlewaite, A., and Samuelson, L. (2014). Stable matching with incomplete information. *Econometrica*, 82(2):541–587.

Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica: Journal of the Econometric Society*, pages 1029–1050.

Pomatto, L. (2022). Stable matching under forward-induction reasoning. *Theoretical Economics*, 17(4):1619–1649.

Shapley, L. S. and Shubik, M. (1971). The assignment game I: The core. *International Journal of game theory*, 1(1):111–130.

Wilson, R. (1978). Information, efficiency, and the core of an economy. *Econometrica: Journal of the Econometric Society*, pages 807–816.

# Appendix

## 7.1 Proof of Proposition 2

**Proposition 2.** Let $V$ be an $n \times n$ surplus matrix. If $\mu \in Ef(V)$, then the above procedure terminates in at most $n$ iteration.

**Proof** : Let $V$ be an $n \times n$ surplus matrix with $\nu_{ij}$ denoting the surplus of the pair $(i, j)$, and $\mu \in Ef(V)$, with $\mu(i) = i$. Since $\mu \in Ef(V)$, then $\sum_{i=1}^{n} \nu_{i\mu(i)} \geq \sum_{i=1}^{n} \nu_{i\mu'(i)}$ for any other assignment $\mu' \neq \mu$. By the procedure, we have $\hat{f}_i^{(\mu,\Sigma^0)} = \nu_{ii}$ for all $i \in N$, and $A^1$:

$$A^1 = \begin{bmatrix} 0 & \max(\nu_{12} - \nu_{22}, 0) & \dots & \max(\nu_{1n} - \nu_{nn}, 0) \\ \max(\nu_{21} - \nu_{11}, 0) & 0 & \dots & \max(\nu_{2n} - \nu_{nn}, 0) \\ \vdots & \vdots & \vdots & \vdots \\ \max(\nu_{k1} - \nu_{11}, 0) & \max(\nu_{k2} - \nu_{22}, 0) & \dots & \max(\nu_{kn} - \nu_{nn}, 0) \\ \vdots & \vdots & \vdots & \vdots \\ \max(\nu_{n1} - \nu_{11}, 0) & \max(\nu_{n2} - \nu_{22}, 0) & \dots & 0 \end{bmatrix}$$

We show that in $a^1$ there is at least one worker that must receive a payoff of 0. On the contrary assume that all $n$ entries in $a^1$ are strictly positive. That is let $a^1$ be as follows,

$$a^1 = \begin{bmatrix} \nu_{1j_1} - \nu_{j_1j_1} \\ \nu_{2j_2} - \nu_{j_2j_2} \\ \vdots \\ \nu_{nj_n} - \nu_{j_nj_n} \end{bmatrix}$$

where $j_i = \underset{j}{\operatorname{argmax}}\, a^1_{ij}$. That is $j_i$ is the column which corresponds to the maximum entry of the $A^1$ matrix in row $i$.

Next we construct a directed graph $G = \langle \Lambda, E \rangle$, with the set of vertices as $\Lambda = \{1, \ldots, n\}$, and there is an edge from $k$ to $j_k$ whenever the $k^{th}$ row of $a^1$ equals $\nu_{kj_k} - \nu_{j_kj_k}$. Note that the graph $G$ has $n$ vertices and $n$ edges, and every vertex has an out degree of 1 (since the out degrees corresponds to the rows of the $a^1$ matrix). Therefore $G$ contains a cycle. Let the cycle be as $\langle (i_1, j_{i_1}), (i_2, j_{i_2}), \ldots (i_t, j_{i_t}) \rangle$. Since, we assumed that every entry of $a^1$ is strictly positive, then we have $\nu_{i_1 j_{i_1}} > \nu_{j_{i_1} j_{i_1}}, \ldots, \nu_{i_t j_{i_t}} > \nu_{j_{i_t} j_{i_t}}$, which implies

$$\sum_{k=1}^{t} \nu_{i_k j_{i_k}} > \sum_{k=1}^{t} \nu_{j_{i_k} j_{i_k}} \tag{9}$$

Consider the set of agents $\Gamma = \{s \mid s \neq i_l, 1 \leq l \leq t\}$. That is $\Gamma$ is the set of agents that do not show in the cycle. Adding the value of $\nu_{ss}$ for all the agents in the set $\Gamma$ to both side of the Equation 9, we have:

$$\sum_{k=1}^{n} \nu_{i_k j_{i_k}} > \sum_{k=1}^{n} \nu_{j_{i_k} j_{i_k}}$$

The right-hand side of the above equation, corresponds to the assignment $j_{i_k} = \mu(j_{i_k})$ for all $k \in \{1, \ldots, n\}$. The left-hand side corresponds to some other assignment $i_k = \mu'(j_{i_k})$. But this contradicts the efficiency of $\mu$. Therefore, at least one entry of the $a^1$ must be 0.

Next we show that there exist a worker and a firm that cannot block with any other worker or firm after the first step. Assume that in $a^1$, there are $t$ rows that are 0. Without loss of generality, we assume that the last $t$ rows are 0. That is $a^1_i = 0$ for all $i \in \{n-t+1, \ldots, n\}$.

Note that, $a^1$ determines the minimum that each worker should get so that the matching $(\mu, \Sigma^0)$ is not blocked. Hence we subtract these values form the initial matrix $V$, to get another surplus matrix $V'$. That is, $V'$ is an $n \times n$ matrix with the $(i,j)^{th}$ entry defined as $\nu'_{ij} = \nu_{ij} - a^1_i$.

$$V' = \begin{bmatrix} \nu_{11} - \nu_{1j_1} + \nu_{j_1j_1} & \nu_{12} - \nu_{1j_1} + \nu_{j_1j_1} & \cdots & \nu_{1n} - \nu_{1j_1} + \nu_{j_1j_1} \\ \nu_{21} - \nu_{2j_2} + \nu_{j_2j_2} & \nu_{22} - \nu_{2j_2} + \nu_{j_2j_2} & \cdots & \nu_{2n} - \nu_{2j_2} + \nu_{j_2j_2} \\ \vdots & \vdots & \vdots & \vdots \\ \nu_{n-t1} - \nu_{n-tj_{n-t}} + \nu_{j_{n-t}j_{n-t}} & \nu_{n-t2} - \nu_{n-tj_{n-t}} + \nu_{j_{n-t}j_{n-t}} & \cdots & \nu_{n-tn} - \nu_{n-tj_{n-t}} + \nu_{j_{n-t}j_{n-t}} \\ \nu_{(n-t+1)1} & \nu_{(n-t+1)2} & \cdots & \nu_{(n-t+1)n} \\ \vdots & \vdots & \vdots & \vdots \\ \nu_{n1} & \nu_{n2} & \cdots & \nu_{nn} \end{bmatrix}$$

Next we, use the iterative elimination procedure on $V'$ to construct the $A'$ matrix and we show that every entry on the last $t$ columns of the $A'$ are always non-positive.

**Claim.** Every entry on the last $t$ columns of the $A'$ are always non-positive.

**Proof** : Consider any column $k \in \{n-t+1, \dots, n\}$.

- For $i \in \{1, \dots, k-1\}$, the $(i,k)^{th}$ entry of $A'$ is $a'^1_{ik} = \max(\nu_{ik} - \nu_{ij_i} + \nu_{j_ij_i} - \nu_{kk}, 0)$. Since for the row $i$ in $A^1$, the maximum came from the $(i, j_i)$ entry then it must be that this entry is larger or equal to that of the $(i,k)$ entry. That is, $\nu_{ij_i} - \nu_{j_ij_i} \geq \nu_{ik} - \nu_{kk}$. Hence, $\nu_{ik} - \nu_{ij_i} + \nu_{j_ij_i} - \nu_{kk} \leq 0$. Therefore, $\max(\nu_{in} - \nu_{ij_i} + \nu_{j_ij_i} - \nu_{nn}, 0) = 0$. This implies that all the entries above the main diagonal of the $k^{th}$ column of $A'$ are zero.

- For $i = k$, it follows from the definition of $a'^1_{kk} = 0$.

- For $i \in \{k+1, \dots, n\}$, the $(i,k)^{th}$ entry of $A'$ is $a'^1_{ik} = \max(\nu_{ik} - \nu_{kk}, 0)$. Since we assumed that the last $t$ rows of $a^1$ are zero, then $\max(\nu_{ik} - \nu_{kk}, 0) = 0$. This implies that all the entries below the main diagonal of the $k^{th}$ column of $A'$ are zero.

■

Next we show that, for least one of the rows in the last $t$ rows of the $A'$ matrix every entry must be zero.

**Claim.** There exist at least one row in the last $t$ rows of $A'$ where every entry is zero.

**Proof** : Take $i \in \{n-t+1, \dots, n\}$. Note that $a'^1_{ik} = 0$, for all $k \in \{n-t+1, \dots, n\}$ by the previous claim. Hence we only have to consider $a'^1_{ik}$, for $k \in \{1, \dots, n-t\}$. That is we have to show the following statement is true.

$$\left(a'^1_{(n-t+1)1} \leq 0 \wedge a'^1_{(n-t+1)2} \leq 0 \wedge \cdots \wedge a'^1_{(n-t+1)(n-t)} \leq 0\right) \bigvee \cdots \bigvee \left(a'^1_{n1} \leq 0 \wedge a'^1_{n2} \leq 0 \wedge \cdots \wedge a'^1_{n(n-t)} \leq 0\right)$$

On the contrary assume the above statement is not true, therefore we have to show that

$$\left(a'^{1}_{(n-t+1)1} > 0 \vee a'^{1}_{(n-t+1)2} > 0 \vee \cdots \vee a'^{1}_{(n-t+1)(n-t)} > 0\right) \bigwedge \cdots \bigwedge \left(a'^{1}_{n1} > 0 \vee a'^{1}_{n2} > 0 \vee \cdots \vee a'^{1}_{n(n-t)} > 0\right)$$

is false. The above statement is equivalent to showing that

$$\left(a'^{1}_{(n-t+1)1} > 0 \wedge a'^{1}_{(n-t+2)1} > 0 \wedge \cdots \wedge a'^{1}_{n1} > 0\right) \bigvee \cdots \bigvee \left(a'^{1}_{(n-t+1)(n-t)} > 0 \wedge a'^{1}_{(n-t+2)(n-t)} > 0 \wedge \cdots \wedge a'^{1}_{n(n-t)} > 0\right) \tag{10}$$

is false. Consider the first term of the above statement we have,

$$\nu_{(n-t+1)1} - \nu_{11} + \nu_{1j_1} - \nu_{j_1j_1} > 0 \wedge \cdots \wedge \nu_{n1} - \nu_{11} + \nu_{1j_1} - \nu_{j_1j_1} > 0 \tag{11}$$

Note that, we assumed that the first $n-t$ rows of the $a^1$ are strictly positive, therefore for any $i \in \{1, \dots n-t\}$,

$$\nu_{ij_i} > \nu_{j_ij_i} \tag{12}$$

Next we construct a directed graph $G = \langle \Lambda, E \rangle$, with the set of vertices as $\Lambda$ indexed by $\{1, \dots, n\}$, and there is an edge from $i$ to $j_i$ whenever $\nu_{ij_i} - \nu_{j_ij_i}$ for $i \in \{1, \dots, n-t\}$. Note that, this graph has no cycle as we showed previously. Next we add an edge for each term in Equation 11. That is we add an edge from $n-t+i$ to 1 for all $i \in \{1, \dots, t\}$. Since $G$ has $n$ vertices and $n$ edges and the out degree of every vertex is 1, then $G$ must contain a cycle that two of it's edges corresponds to one of the terms of the form $\nu_{(n-t+i)1} + \nu_{1j_1}$. Let the cycle be as $\langle (i_1, j_{i_1}), (i_2, j_{i_2}), \dots (i_t, j_{i_t}) \rangle$. If the edge is due to Equation 12, then we have $\nu_{ij_i} > \nu_{j_ij_i}$, and if the edge corresponds to one of the terms in Equation 11, then we have $\nu_{(n-t+i)1} + \nu_{1j_1} > \nu_{11} + \nu_{j_1j_1}$. This implies

$$\sum_{k=1}^{t} \nu_{i_kj_{i_k}} > \sum_{k=1}^{t} \nu_{j_{i_k}j_{i_k}} \tag{13}$$

Consider the set of agents $\Gamma = \{s \neq i_1, i_2, \dots, i_t\}$. That is $\Gamma$ is the set of agents that do not show in the cycle. Adding the value of $\nu_{ss}$ for all the agents in the set $\Gamma$ to both side of the Equation 13, we have:

$$\sum_{k=1}^{n} \nu_{i_kj_{i_k}} > \sum_{k=1}^{n} \nu_{j_{i_k}j_{i_k}}$$

The right-hand side of the above equation, corresponds to the assignment $j_{i_k} = \mu(j_{i_k})$ for all $k \in \{1, \dots, n\}$. The left-hand side corresponds to some other assignment $i_k = \mu'(j_{i_k})$. But this contradicts the efficiency of $\mu$. A similar argument shows that every term in 10 cannot be true, hence the entire statement is false, which shows that there is at least a row with all the entries being zero. ■

Let $s$ be the row for which the corresponding row and column in $A'$ is zero (in case there are multiple such rows the following same argument applies). Note that, by construction we have $a^2 = a'^1 + a^1$. Therefore, the row $s$ will remain unchanged after the first iteration. Therefore, we can repeat the procedure by removing the row and column $s$ from the $V'$ matrix, which implies a matrix of size $n-1 \times n-1$. Hence, repeating the argument shows that it takes at most $n$ iteration for the procedure to terminate, as in every iteration one pair of agents can be eliminated from the surplus matrix. ■

## 7.2 Proof of Proposition 5

**Proposition 5**. Suppose $\mu \notin Ef(V)$, then there exists no wage vector $p$ such that $(\mu, p) \in \Sigma^*$.

**Proof** : In what follows we show that, given that the assignment $\mu$ is not efficient, then by the restricted iterated elimination procedure, there will exist some worker such the minimum payoff that she should receive must be larger than the available surplus in the assignment $\mu$. Therefore, by applying Lemma 1, the assignment $\mu$ can not be supported by any payment vector.

Formally, for assignment $\mu$, we define a sequence of recursive functions $\mu^k : I \to I$ for all $k \geq 1$ and $\rho^k : I \to I$ for all $k \geq 0$ as follows:

$$\mu^1(i) = \mu(i) \qquad\qquad \rho^0(i) = i$$
$$\mu^k(i) = \mu(\mu^{k-1}(i)) \qquad\qquad \rho^k(i) = \mu^{-1}(\rho^{k-1}(i))$$

For assignment $\mu \notin Ef(V)$, we show that for some iteration $k$ and some worker $i$, the minimum payoff of worker $i$ exceeds the maximum surplus available to $(i, \mu(i))$. This together with Lemma 1 completes the proof.

After the first iterative elimination of blocked matching outcomes, there exists at least one worker $i$ such that the minimum payoff that worker $i$ receives in assignment $\mu$ is strictly greater than 0, i.e., $\tilde{a}_i^1 > 0$. To show this, on the contrary assume for every worker $i \in I$,

$$\tilde{a}_i^1 = \nu_{ii} - \nu_{\rho^1(i)i} \leq 0$$

Since by definition $\rho^1(i) = \mu^{-1}(i)$, then summing over all $i \in I$, we have $\sum_{i=1}^{n} \nu_{ii} \leq \sum_{i=1}^{n} \nu_{\mu^{-1}(i)i}$, which contradicts with Equation 8.

Next we consider two possible cases:

**Case 1.** After the first iterative elimination of blocked matching outcomes, all workers receive a minimum payoff strictly greater than 0, i.e., $\tilde{a}_i^1 > 0$ for all workers $i$.

Consider worker with index $i$. Let $\mu^m(i) = i$ where $m \leq n$. As $\tilde{a}_i^1 > 0$, worker with index $\mu(i)$ infers that the maximum the firm $\mu(i)$, can earn in assignment $\mu$ is reduced by $\tilde{a}_i^1$. The minimum that worker $\mu(i)$ should receive for the assignment $\mu$ to be not blocked after the second iteration increases to $\tilde{a}^2_{\mu(i)}$ which is strictly greater than $\tilde{a}^1_{\mu(i)}$. Worker with index $\mu^2(i)$ infers that the maximum the firm $\mu^2(i)$, can earn in assignment $\mu$ is reduced by $\tilde{a}^2_{\mu^2(i)}$. The minimum that worker with index $\mu^3(i)$ should receive for the assignment $\mu$ to be not blocked after the third iteration increases to $\tilde{a}^3_{\mu^3(i)}$ which is strictly greater than $\tilde{a}^1_{\mu^3(i)}$. This process continues, when in iteration $(m-1)$, worker with index $i$ infers that the maximum firm $\mu^m(i) = i$, can earn in assignment $\mu$ is reduced by $\tilde{a}^{(m-1)}_{\mu^{m-1}(i)}$. The minimum that worker with index $i$ should receive for the assignment $\mu$ to be not blocked after the $m^{th}$ iteration increases to $\tilde{a}_i^m$ which is strictly greater than $\tilde{a}_i^1$. Worker with index $\mu(i)$ infers that the maximum firm $\mu(i)$ can receive in assignment $\mu$ is further reduced. The minimum that worker with index $\mu(i)$ should receive for the assignment $\mu$ to be not blocked after the $(m+1)^{th}$ iteration increases to $\tilde{a}^{m+1}_{\mu(i)}$ which is strictly greater than $\tilde{a}^2_{\mu(i)}$. The iterative process continues until there is some iteration $k$ and worker $i$ with $\tilde{a}_i^k > \nu_{i\mu(i)}$.

**Case 2.** There exists worker $i$ such that $\tilde{a}_i^1 \leq 0$.

Let $S^1$ be the set of workers such that for every worker $i \in S^1$, $\tilde{a}_i^1 > 0$. If there exists $i \in S^1$ such that $\tilde{a}_i^1 > \nu_{i\mu(i)}$, then using Lemma (1) the proof is completed. Suppose not, consider the following partition of the set of workers $I \setminus S^1$. Let $A^0 \subseteq I \setminus S^1$ be the set of workers $i$ such that $\tilde{a}^1_{\rho^1(i)} \leq 0$. In general let $A^k \subseteq I \setminus S^1$ be the set of workers $i$ such that

$$\begin{aligned} &\tilde{a}^1_{\rho^l(i)} > 0 \Rightarrow \nu_{\rho^l(i)\rho^l(i)} - \nu_{\rho^{l+1}(i)i} > 0 \qquad \text{for} \quad 1 \leq l \leq k \\ \text{and} \quad &\tilde{a}^1_{\rho^{k+1}(i)} \leq 0 \end{aligned}$$

We partition the set of workers $I \backslash S^1$ into collection of sets $\{A^0, A^1, \ldots, A^{n-1}\}$. Consider a partition element $A_k$ where $k \geq 1$ and some worker $i$. In assignment $\mu$, worker with index $\rho^k(i)$ is matched to firm with index $\rho^{k-1}(i)$ and $\tilde{a}^1_{\rho^k(i)} > 0$. Worker with index $\rho^{k-1}(i)$ infers that the maximum firm with index $\rho^{k-1}(i)$ can receive in assignment $\mu$ is reduced.

The minimum that worker with index $\rho^{k-1}(i)$ should receive for the assignment $\mu$ to be not blocked after the second iteration increases to $\tilde{a}^2_{\rho^{k-1}(i)}$ which is strictly greater than $\tilde{a}^1_{\rho^{k-1}(i)}$. As in assignment $\mu$ worker with index $\rho^{k-1}(i)$ is matched to firm with index $\rho^{k-2}(i)$ and $\tilde{a}^2_{\rho^{k-1}(i)} > 0$. Worker with index $\rho^{k-2}(i)$ infers that the maximum firm with index $\rho^{k-2}(i)$ can receive in assignment $\mu$ is reduced. The minimum that worker with index $\rho^{k-2}(i)$ should receive for the assignment $\mu$ to be not blocked after the third iteration increases to $\tilde{a}^3_{\rho^{k-1}(i)}$ which is strictly greater than $\tilde{a}^1_{\rho^{k-2}(i)}$. This process continues until worker with index $i$ infers that the maximum firm with index $\rho^1(i)$ can receive in assignment $\mu$ is reduced. The minimum worker with index $i$ should receive for the assignment $\mu$ to be not blocked after the $k^{th}$ iteration is $\max(\tilde{a}^k_i, 0)$ where

$$\tilde{a}^k_i = \sum_{l=0}^{k} \left( \nu_{\rho^l(i)\rho^l(i)} - \nu_{\rho^{l+1}(i)\rho^l(i)} \right)$$

We now show that there exists a partition element $A_k \in A$ and worker $i \in A_k$ such that $\tilde{a}^k_i > 0$. Suppose not, that is for every partition element $A_k$ and for every worker $i \in A_k$, $\tilde{a}^k_i \leq 0$, i.e.,

$$\sum_{l=0}^{k} \left( \nu_{\rho^l(i)\rho^l(i)} - \nu_{\rho^{l+1}(i)\rho^l(i)} \right) \leq 0$$

Summing over all workers $i$ in $A_k$ and for every partition element $A_j$,

$$\sum_{A_k} \sum_{i \in A_k} \sum_{l=0}^{k} \left( \nu_{\rho^l(i)\rho^l(i)} - \nu_{\rho^{l+1}(i)\rho^l(i)} \right) = \sum_{i=1}^{n} \left( v_{ii} - v_{\mu^{-1}(i)i} \right) \leq 0$$

However, this contradicts with Equation 8.

Hence, there exists a subset of workers $T^1 \subseteq I \setminus S_1$ such that for $i \in T^1$, $\tilde{a}^k_i > 0$ where $i \in A_k$. Let $k_1$ denote the first (minimum) iteration such that $\tilde{a}^{k_1}_i > 0$ for all $i \in T^1$. Let $S^2$ be the set of workers such that $\tilde{a}^{k_1}_i > 0$ for every worker $i \in S^2$. Then,

$$S^2 = S^1 \cup T^1$$

If there exists $i \in S^2$ such that $\tilde{a}^{k_1}_i > \nu_{i\mu(i)}$, then using Lemma 1, completes the proof. If not, then again we consider a similar partition of the set of workers $I \setminus S^2$ into collection of sets $\{A^1, A^2, \ldots, A^{n-1}\}$ and show that there exists a subset of workers $T^2 \subseteq I \setminus S^2$ such that for $i \in T^2$, $\tilde{a}^{(k_1+k_2)}_i > 0$ where $i \in A_l$.

For, $3 \leq m \leq n-1$, let $S^m$ be the set of workers such that $\tilde{a}_i^{\left(\sum\limits_{i=1}^{m-1} k_i\right)} > 0$, for all $i \in S^m$.

$$S^m = S^{m-1} \cup T^{m-1}$$

If there exists $i \in S^m$ such that $\tilde{a}_i^{\left(\sum\limits_{i=1}^{m-1} k_i\right)} > \nu_{i\mu(i)}$, then again using Lemma 1, completes the proof.

If for some $m \leq (n-1)$, $|S^m| = n-1$, let worker $\{j\} = I \setminus S^m$, then

$$\tilde{a}_j^{\left(\sum_{i=1}^{m-1} k_i+1\right)} = \sum_{l=0}^{n} \left(\nu_{\rho^l(j)\rho^l(j)} - \nu_{\rho^{l+1}(j)\rho^l(j)}\right) = \epsilon > 0$$

Hence, we get the same condition as in Case 1. Worker with index $\mu(j)$ infers that the maximum firm $\mu(j)$ can earn in assignment $\mu$ is reduced by $\epsilon$. Hence, minimum payoff worker with index $\mu(j)$ should receive increases by $\epsilon$.

$$\tilde{a}_{\mu(j)}^{\left(\sum_{i=1}^{m-1} k_i+2\right)} = \tilde{a}_{\mu(j)}^{\left(\sum_{i=1}^{m-1} k_i+1\right)} + \epsilon$$

Worker with index $\mu(\mu(j))$ infers this and the recursive process continues until there is some worker $i$ and $l \in \mathrm{N}$ such that

$$\tilde{a}_i^{\left(\sum_{i=1}^{m-1} k_i+l\right)} > \nu_{i\mu(i)}$$

Then using Lemma 1, completes the proof.

Similarly, if for some $m \leq (n-1)$, $|S^m| = (n-t)$ for $2 < t \leq (n-1)$ and $|S^{m+1}| = n$, then for every worker $j \in T^m$:

$$\tilde{a}^{\left(\sum_{i=1}^{m-1} k_i+1\right)}(j) > 0$$

Hence, for every worker $j \in T^m$, the maximum firm $\mu(j)$ can earn is now reduced by a positive amount. Worker with index $\mu(\mu(j))$ infers this and the minimum wage that she can earn is increased by that same positive amount. This recursive updating process continues until there is some worker $i$ and $l \in \mathrm{N}$ such that

$$\tilde{a}_i^{\left(\sum_{i=1}^{m-1} k_i+l\right)} > \nu_{i\mu(i)}$$

Then using Lemma 1, completes the proof. ■