von Neumann-Morgenstern stable sets in job matching problems*

Madhuparna Karmokar and Debasis Mishra †

November 23, 2026

Abstract

We show that a von Neumann-Morgenstern (vNM) stable set generically exists in the many-to-one job-matching model. Unlike the core, which can be empty in this model, a vNM stable set exists whenever there is a unique efficient matching: the set of all individually rational (IR) and efficient outcomes form a vNM stable set. If an economy has multiple efficient matchings, a vNM stable set may not exist. If the vNM stable sets are required to be *rich* in a certain sense, then such a rich vNM stable set exists if and only if the set of efficient matchings is *essentially* unique, and the set of all efficient and IR outcomes forms the unique rich vNM stable set.

Keywords: job-matching, stable sets, core, individually rational outcomes.

JEL Classification: D82, L51

*We are grateful to Jean-Jacques Herings and participants at a workshop on matching at Indian Statistical Institute, Delhi for comments and feedback. `refine.ink` was used to check consistency of proofs and exposition of this article. Debasis Mishra acknowledges financial support from ANRF through Project ANRF/ARGM/2025/000572/QSS.

†Karmokar: Indian Institute of Management Calcutta, `madhuparnak@iimcal.ac.in`. Mishra: Indian Statistical Institute, Delhi, `dmishra@isid.ac.in`.

1 Introduction

In many labor markets, firms hire teams of workers, and the surplus generated by a team of workers exhibit complementarity. Such complementarities lead to settings where a stable outcome may not exist, a fundamental desiderata in matching theory. In the many-to-one job-matching model, stability was first studied by Kelso and Crawford (1982), who showed the existence of a stable outcome when utility functions of firms over workers satisfy the gross substitutes condition – see Hatfield et al. (2019); Baldwin and Klemperer (2019); Kojima et al. (2020) for recent generalizations of this result.[1]

In this paper, we seek an alternate solution: the vNM stable set, which was introduced as a set-valued solution concept in the magnum opus von Neumann and Morgenstern (1944). Unlike core, where each outcome in the core is stable, individual outcomes in a vNM stable set have no meaning. It is the entire set that is stable. This is the reason von Neumann and Morgenstern (1944) interpret stable sets as "standards of behavior". A vNM stable set is a set of individually rational (IR) outcomes (an outcome is a matching and payoffs to firms and workers), which are required to be consistent using a domination relation over the set of all IR outcomes.[2] vNM stability requires two consistency properties: (a) *internal stability* – outcomes inside the vNM stable set do not dominate each other; (b) *external stability* – each IR outcome not in the vNM stable set is dominated by some outcome inside.

In the job-matching model, a vNM stable set could be interpreted as a menu of IR outcomes/contracts (matching and non-negative payoff splits) that firms and workers may choose from. This menu does not undermine itself (internal stability) and anything not on the menu has a counter-contract on the menu. Hence, anything not in the menu cannot be included because it will be immediately countered by something already in the menu (external stability). Further, anything in the menu cannot be excluded because nothing already in the menu can counter it (internal stability). Hence, a vNM stable menu of contracts cannot exclude any of its members without losing its justification for exclusion,

[1] There have been other approaches to show existence of stable matching beyond substitutable preferences of firms. Che et al. (2019) shows that in a large market with a continuum of workers a stable matching exists even when firms' preference exhibit complementarities. Rostek and Yorder (2020) establish the existence of stable outcomes in contracting environments with complementarities, allowing for multilateral agreements and externalities. In their transferable-utility setting, primitive contracts satisfy gross complementarity for every agent. These conditions do not generally hold in our many-to-one environment, where a worker can be employed by at most one firm and firms' surplus functions are unrestricted.

[2] According to this domination relation, outcome a dominates outcome b, if there exists a firm and its matched workers in a get higher utility in a than in b (with one agent strictly higher). This is akin to blocking the outcome b in the sense of Kelso and Crawford (1982).

and cannot admit any outsider without introducing an internal conflict.

Our main result shows that a vNM stable set in the job-matching model exists whenever the efficient matching is unique. The set of all IR and efficient outcomes forms a vNM stable set in that case. Thus, even when teams of workers exhibit significant complementarities, a vNM stable set exists, though a stable outcome may not exist. We show that if there are multiple efficient matchings, then (a) the vNM stable set may not exist; and (b) even when it exists, it may contain inefficient outcomes. However, by requiring the vNM stable set to be rich in a certain sense, we establish that the set of IR and efficient outcomes is essentially the unique vNM stable set. Hence, a rich vNM stable set mandates firms and workers to agree on a menu of contracts that contains only efficient matchings but leaves the division of surplus completely flexible.

Our richness is the following requirement. Call a vNM stable set $\mathcal{V}$ *firm-rich* if for every outcome (with matching, say, μ) in $\mathcal{V}$, another outcome with the same matching μ and that assigns zero payoffs to all workers belongs to $\mathcal{V}$ – such an outcome is the firm-optimal IR outcome corresponding to μ.[3] We show that in a generic economy, a firm-rich vNM stable set exists if and only if there is a unique efficient matching.[4] In that case, the set of all IR and efficient outcomes forms the unique firm-rich vNM stable set. Firm-richness requires the extreme outcome, firm-IR optimal outcome, to be included in the vNM stable set. This is plausible in settings where firms have a lot of bargaining power in deciding the menu of contracts. If the economy is non-generic, a similar result holds under more richness, where the extreme outcomes giving each individual worker all the surplus is included in the menu of outcomes. For such rich vNM stable sets, we not only need the firm-IR optimal outcome but also *worker-IR maximal* outcomes to be included in the vNM stable sets.

Unlike vNM stable sets, rich vNM stable sets exhibit simpler structure, and are easier to analyze. This allows us to make progress toward a characterization of rich vNM stable sets. In general, vNM stable sets are known to be difficult to analyze. Quoting Aumann (1985)

> *What has prevented the (vNM stable set) theory from being more successful is that it is so difficult to work with. Finding stable sets involves a new tour de force of mathematical reasoning for each game or class of games that is considered.*

[3] Given a matching μ, the firm-IR optimal outcome assigns zero payoffs to all workers and gives the entire surplus from μ to firms.

[4] The genericity requirement is a no-tie condition on the surplus generated by matching firms to teams of workers. In particular, it says that every firm generates different surplus when matched to different teams of workers.

Our results contrast with the results in the matching literature without transfer, where a unique vNM stable set exists in the one-to-one matching model (Ehlers, 2007; Wako, 2010).[5]

The job-matching model induces a transferrable utility (TU) cooperative game. The study of vNM stable sets in general cooperative games has a long history – see Aumann (1985) for a very early survey and Subiza et al. (2025) for a recent contribution. This literature defines outcomes to be imputations (IR and efficient outcomes) of the cooperative game. In the job-matching model, this will loosely translate to finding a vNM stable set in the set of efficient and individually rational outcomes. This approach is taken in Núñez and Rafels (2013), who analyze the assignment game (one-to-one job-matching model of Shapley and Shubik (1971)) and identify vNM stable sets. We allow for inefficient outcomes to be included in the vNM stable sets. Indeed, we show that vNM stable sets may contain inefficient outcomes. However, our partial converse result establishes that we get efficiency as a consequence of rich vNM stability.[6]

The domination relation used to define the vNM stable set reflects a myopic notion of blocking in the job-matching problem. An interesting direction of further research is to consider a farsighted notion of blocking. Farsighted notions of blocking and related vNM stable sets have been studied in matching literature without transfers by Mauleon et al. (2011); Klaus et al. (2011); Herings et al. (2020); Doğan and Ehlers (2026). The nature of dominance relation used to define vNM farsighted stable set considers a sequence of blockings, and has different implications on the structure of stable sets. In the future, we plan to investigate the structure of vNM farsighted stable set of the job-matching problem.

Set-valued stability notions have been recently used in matching with incomplete information. In an important paper, Liu et al. (2014) introduce the notion of a self-stabilizing set and show its existence in the one-to-one matching model of Shapley and Shubik (1971) with one-sided incomplete information. Their notion is similar to the rationalizability (iterated elimination) idea in game theory but applied to their model with incomplete information. Liu (2024) is a recent survey of this growing literature. Pourpouneh et al. (2026) study a similar set-valued concept of stability in the one-to-one matching model of Shapley and Shubik (1971) where agents do not observe the outcomes of the other agents. They show that a subset of efficient and IR outcomes containing the core forms the largest self-stabilizing set.

[5]Other papers studying vNM stable sets in one-to-one matching model (without transfers) include Herings et al. (2017); Doğan and Ehlers (2026).

[6]There are other differences between our approach and the TU game approach to vNM stability – see Remark 1.

While this literature and our paper look at set-valued stability solutions, there are significant differences. First, as emphasized earlier, individual outcomes have no meaning in our solution and it is only the collection that is stable. This is not the case in this literature. Second, the set-valued stability is driven by incomplete information in these papers, but we have complete information. Third, the main method to reach a stable set in these papers is an iterated elimination procedure using the uncertainty in these models. This is not the case for us.

The paper is organized as follows. Section 2 introduces the job-matching model. Section 3 defines stability, a domination relation among IR outcomes, and vNM stable set. Section 4 proves our first result on existence of vNM stable sets. Section 5 introduces rich vNM stable sets and proves our results on uniqueness of rich vNM stable sets. Appendix A provides a detailed proof of these results, but all other omitted proofs are in Appendix B. We conclude in Section 6 with a discussion on our results.

2 Model

We study the standard model of job-matching due to Kelso and Crawford (1982). Let F be the set of firms and W be the set of workers. A matching is a map $\mu : W \to F \cup \{\emptyset\}$, where $\mu(w)$ denotes the firm worker w is matched to – if $\mu(w) = \emptyset$, then it denotes an unmatched worker. For every $f \in F$, we denote by $\mu(f) := \{w : \mu(w) = f\}$ the set of workers matched to firm f in μ.

For every $f \in F$ and every $S \subset W$, let $V_f(S)$ denote the **surplus** generated by matching workers in S to firm f. The surplus function V_f may be generated from value function $v_f : 2^W \to \Re_+$ of the firm (with $v_f(\emptyset) = 0$) and disutility function $v_w : F \cup \{\emptyset\} \to \Re_+$ of each worker w:

$$V_f(S) = v_f(S) - \sum_{w \in S} v_w(f)$$

where $v_w(\emptyset) = 0$. Combined with salaries, this generates utilities for firms and workers as in Kelso and Crawford (1982). However, the surplus function V_f of each firm f is a sufficient reduced form for our purposes. As the surplus function is generated from value of firm and disutilities of workers, (a) it can take negative values; (b) it need not be monotone.

An **outcome** is a pair (μ, u), where μ is a matching and $u : W \to \Re$ corresponds to the

utility assigned to *workers*. An outcome specifies the utility assigned to each firm indirectly. In particular, in outcome (μ, u), each firm f receives utility equal to

$$V_f(\mu(f)) - \sum_{w \in \mu(f)} u(w)$$

Such a division of surplus is possible because utility is transferrable (using salaries) in this model. This also explicitly assumes budget-balance among a firm f and its matched workers $\mu(f)$. That is, transferring surplus across firms is not possible in our model. This is consistent with much of the literature on job-matching, and makes sense because firms are independent financial entities. Also, note that by our definition of an outcome, any worker who is unmatched must get zero utility.

An economy consists of the set of firms, the set of workers, and the surplus functions:

$$\Big(F, W, \{V_f\}_{f \in F}\Big)$$

An outcome (μ, u) is **individually rational (IR)** if

$$\begin{aligned} u(w) &\geq 0 \qquad && \forall\ w \in W \\ V_f(\mu(f)) - \sum_{w \in \mu(f)} u(w) &\geq 0 \qquad && \forall\ f \in F, \end{aligned}$$

As we emphasized earlier, by the definition of an outcome, the first inequality is an equality for any worker who is unmatched.

The two constraints correspond to IR of workers and firms respectively. Let $\mathcal{I}$ be the set of all IR outcomes.

A matching $\mu^\star$ is **efficient** if

$$\mu^\star \in \arg\max_{\mu} \sum_{f \in F} V_f(\mu(f))$$

Clearly, in an efficient matching $V_f(\mu^\star(f)) \geq 0$ for all firms $f \in F$. An outcome (μ, u) is efficient if μ is efficient. Notice that the only way an outcome is *inefficient* if the matching is not efficient. This is because our requirement of budget balance inside a single firm – any outcome must split the surplus raised by a matched firm and workers between them.

3 Stable outcomes, domination, and vNM stable sets

In this section, we define the notion of vNM stable sets. It will require us to define a relation over the set of IR outcomes. To do so, it is useful to remind the notion of blocking and stability first.

3.1 Stable outcomes

A firm $f \in F$ and a subset of workers $S \subseteq W$ **block** an outcome (μ, u) if there exists a utility vector $\hat{u} : S \to \Re_+$ such that

$$\begin{aligned} \hat{u}(w) &\geq u(w) \qquad \forall\ w \in S \\ V_f(S) - \sum_{w \in S} \hat{u}(w) &\geq V_f(\mu(f)) - \sum_{w \in \mu(f)} u(w) \end{aligned}$$

with at least one inequality strict. In other words, firm f and workers in S can be better off sharing the surplus $V_f(S)$ using $\hat{u}$ than the current outcome (μ, u). An outcome $(\mu, u) \in \mathcal{I}$ is **stable** if it is not blocked by any (f, S) pair where $f \in F$ and $S \subseteq W$ and it is individually rational. Notice that $(f, \mu(f))$ cannot block an outcome (μ, u) – putting $S = \mu(f)$ in the blocking inequalities above and adding them gives $0 > 0$. Also, it is well known that a stable outcome may not exist, and the set of stable outcomes is also the *core* of the induced transferrable utility cooperative game.

3.2 A domination relation among IR outcomes

The following relation among outcomes in $\mathcal{I}$ is used to define vNM stability.

Definition 1. *An outcome $(\mu, u) \in \mathcal{I}$ is* **dominated** *by an outcome $(\mu', u') \in \mathcal{I}$, written as $(\mu', u') \to (\mu, u)$, if there exists a firm f such that*

$$\begin{aligned} u'(w) &\geq u(w) \qquad \forall\ w \in \mu'(f) \\ V_f(\mu'(f)) - \sum_{w \in \mu'(f)} u'(w) &\geq V_f(\mu(f)) - \sum_{w \in \mu(f)} u(w) \end{aligned}$$

with one inequality strict. We denote this as $(\mu', u') \xrightarrow{f} (\mu, u)$.

Hence, $(\mu', u') \to (\mu, u)$ requires that some firm f and workers $\mu'(f)$, who are matched to f in μ', block (μ, u) (using u'). Since it is possible that two different firms evaluate these two outcomes differently, the relation $\to$ is not necessarily anti-symmetric. It also does not define a complete relation on $\mathcal{I}$. If two outcomes (μ, u) and (μ', u') are not comparable under $\to$, we write $(\mu, u) \nleftrightarrow (\mu', u')$.

For every matching μ, let $\mathcal{I}_\mu$ denote the set of all IR outcomes associated with μ:

$$\mathcal{I}_\mu := \{(\mu, u) : (\mu, u) \in \mathcal{I}\}$$

Since surplus function V_f of each firm f need not be non-negative, $\mathcal{I}_\mu$ may be empty. For instance, consider an economy with only one firm and one worker and suppose the disutility of worker is larger than the value it generates for the firm, i.e., the surplus is negative. Then the matching μ that matches the worker to the firm has no IR outcome (μ, u). Hence, $\mathcal{I}_\mu = \emptyset$. Here, the efficient matching is to not match the worker to the firm, in which everyone getting zero payoff is IR. In general, if $\mu^\star$ is an efficient matching, $V_f(\mu^\star(f)) \geq 0$ for all f and hence, the zero payoff vector $\mathbf{0}$ ($u_w = 0$ for all w) generates the IR outcome $(\mu^\star, \mathbf{0})$. Hence, $\mathcal{I}_{\mu^\star}$ is always non-empty.

Two outcomes with the same matching are not comparable according to relation $\to$. This follows immediately because for blocking the matching needs to change. We record this as an observation.

Observation 1. *For every matching μ, $(\mu, u) \nleftrightarrow (\mu, u')$ for all $(\mu, u), (\mu, u') \in \mathcal{I}_\mu$.*

3.3 vNM stable sets

We now define vNM stability using the relation $\to$.

Definition 2. *A set of outcomes $\mathcal{V} \subseteq \mathcal{I}$ is a* **vNM stable set** *if*

- *no outcome in $\mathcal{V}$ is dominated by another outcome in $\mathcal{V}$: for every $(\mu, u), (\mu', u') \in \mathcal{V}$,*

$$(\mu, u) \nleftrightarrow (\mu', u') \qquad \textbf{(Internal stability)}$$

- *every outcome not in $\mathcal{V}$ is dominated by some outcome in $\mathcal{V}$: for every $(\mu, u) \in \mathcal{I} \setminus \mathcal{V}$,*

there exists $(\mu', u') \in \mathcal{V}$ *such that*

$$(\mu', u') \to (\mu, u) \qquad \textbf{(External stability)}$$

Remark 1. We have chosen to consider the set of IR outcomes as the set of feasible outcomes. This set will include inefficient outcomes too. This is a departure from how the TU cooperative game theory literature studies vNM stable sets. In TU games studying vNM stable sets, an outcome is an *imputation* vector of players in the game, where an imputation vector is defined to be efficient (sum of imputations of players gives the value of the grand coalition of the TU game) and IR (Peters, 2008; Núñez and Rafels, 2013; Subiza et al., 2025). We find no reason to restrict attention to efficient outcomes in our model. Indeed, Appendix C contains an example that has a vNM stable set with *only inefficient* outcomes (where core is empty). Further, the study of vNM stable sets in the matching literature considers inefficient matchings – see Ehlers (2007). Also, even our IR and efficient outcomes are different from imputations in TU games because we require budget-balance between every firm and its matched workers – cross-subsidizing payoffs across firms is not allowed in our outcomes, which is consistent with Kelso and Crawford (1982). ■

If the set of stable outcomes is non-empty, it is contained in every vNM stable set. This is a standard fact but we document it as an observation.

Observation 2. *If* $(\mu^\star, u^\star)$ *is a stable outcome, then* $(\mu^\star, u^\star) \in \mathcal{V}$ *for any vNM stable set* $\mathcal{V}$*.*

The proof of this observation follows from external stability: if $(\mu^\star, u^\star) \notin \mathcal{V}$, then external stability will mean some outcome $(\mu, u) \in \mathcal{V}$ will dominate $(\mu^\star, u^\star)$; but that will mean a firm f and the subset of workers $\mu(f)$ will block $(\mu^\star, u^\star)$, a contradiction to stability of $(\mu^\star, u^\star)$.

4 Existence of vNM stable sets

A vNM stable set may not exist as the following example shows.

Example 1 (vNM stable set may not exist)**.** Consider the following economy. There are three firms $F = \{f_1, f_2, f_3\}$ and three workers $W = \{w_1, w_2, w_3\}$. Firms are *single-minded*: firm f_1 can generate positive (equal to 1) surplus with workers $\{w_1, w_2\}$ only; firm f_2 can generate positive (equal to 1) surplus with workers $\{w_2, w_3\}$ only; and firm f_3 can generate

positive (equal to 1) surplus with workers $\{w_3, w_1\}$ only. This is given in Table 1.[7] There

	$\{w_1\}$	$\{w_2\}$	$\{w_3\}$	$\{w_1, w_2\}$	$\{w_2, w_3\}$	$\{w_3, w_1\}$	$\{w_1, w_2, w_3\}$
$V_{f_1}(\cdot)$	-1	-1	0	1	0	0	0
$V_{f_2}(\cdot)$	0	-1	-1	0	1	0	0
$V_{f_3}(\cdot)$	-1	0	-1	0	0	1	0

Table 1: Surplus of firms and workers

are three efficient matchings: μ_1 matches f_1 to workers $\{w_1, w_2\}$; μ_2 matches f_2 to workers $\{w_2, w_3\}$ and μ_3 matches f_3 to workers $\{w_3, w_1\}$. In any other matching the total surplus generated is less than or equal to zero and IR requires that all the firms and workers must get zero payoff in any outcome corresponding to such matchings. So, if $\mu \notin \{\mu_1, \mu_2, \mu_3\}$, then $(\mu, u_{w_1}, u_{w_2}, u_{w_3}) \in \mathcal{I}$ implies $u_{w_k} = 0$ for all $k \in \{1, 2, 3\}$. We show that no vNM stable set exists for this example.

Lemma 1. *A vNM stable set does not exist for the economy in Example 1.*

While vNM stable set does not exist in Example 1, it is a *non-generic* example in the sense that there are multiple efficient matchings. We now show that if a unique efficient matching exists, a vNM stable set always exists: the set of all efficient and IR outcomes is one such set.[8]

Theorem 1. *If there is a unique efficient matching, then the set of all IR and efficient outcomes is a vNM stable set.*

Proof. Internal stability of $\mathcal{I}_{\mu^\star}$ follows from Observation 1. The following lemma, whose proof is given in Appendix B, establishes external stability.

Lemma 2. *Suppose μ is not an efficient matching and $(\mu, u) \in \mathcal{I}$. Then, there exists $(\mu^\star, u^\star) \in \mathcal{I}_{\mu^\star}$ such that $\mu^\star$ is efficient and $(\mu^\star, u^\star) \to (\mu, u)$.*

Since the unique efficient matching is $\mu^\star$, any $(\mu, u) \notin \mathcal{I}_{\mu^\star}$ is inefficient. Hence, Lemma 2 implies there exists $(\mu^\star, u^\star) \in \mathcal{I}_{\mu^\star}$ such that $(\mu^\star, u^\star) \to (\mu, u)$. □

[7] We have chosen negative surplus values for some entries in Table 1. This is only to make the analysis simpler. The example can be modified to make all surplus entries non-negative and still the vNM stable set will be empty.

[8] Fixing the disutilities of workers, the set of utility functions of firms where a unique efficient matching exists has Lebesgue measure one in the set of all utility functions of firms.

STABLE OUTCOMES AND VNM STABLE SETS. By Observation 2, whenever the core is non-empty, it is included in every vNM stable set. We now share two observations about vNM stable sets when core is non-empty: (a) core may not be a vNM stable set; (b) a vNM stable set may contain inefficient outcomes along with core.

Corollary 1. *Suppose $\mu^\star$ is the unique efficient matching and the core $\mathcal{C}$ is non-empty. Then, $\mathcal{C}$ is a vNM stable set if and only if $\mathcal{C} = \mathcal{I}_{\mu^\star}$.*

Proof. The fact that $\mathcal{I}_{\mu^\star}$ is a vNM stable set follows from Theorem 1. For the other direction, suppose $\mathcal{C}$ is a vNM stable set. Since every outcome in $\mathcal{C}$ is efficient, we have $\mathcal{C} \subseteq \mathcal{I}_{\mu^\star}$. Assume for contradiction $\mathcal{C} \subsetneq \mathcal{I}_{\mu^\star}$. Now, pick $(\mu^\star, u^\star) \in \mathcal{I}_{\mu^\star} \setminus \mathcal{C}$. By Observation 1, for every $(\mu^\star, u) \in \mathcal{C}$, we have $(\mu^\star, u) \nrightarrow (\mu^\star, u^\star)$. Hence, $\mathcal{C}$ violates external stability, a contradiction. □

Usually core is a strict subset of the set of IR outcomes corresponding to efficient matchings. Hence, when there is a unique efficient matching, core is not usually vNM stable. The following example shows that even when there are multiple efficient matchings, the core may be a strict subset of a vNM stable set and the vNM stable set may contain *inefficient* matchings.

Example 2 (vNM stable set with inefficient outcomes)**.** Let $F = \{f\}$ and $W = \{w_1, w_2\}$. Let

$$V_f(\{w_1, w_2\}) = 1, V_f(\{w_1\}) = V_f(\{w_2\}) = 2.$$

Consider the four possible matchings:

$$\mu_1 : \mu_1(f) = \{w_1\}; \quad \mu_2 : \mu_2(f) = \{w_2\}; \quad \mu_3 : \mu_3(f) = \{w_1, w_2\}; \quad \mu_4 : \mu_4(f) = \emptyset$$

Out of these, μ_1 and μ_2 are efficient. There are exactly two stable outcomes: $(\mu_1, 0, 0)$ and $(\mu_2, 0, 0)$. So core consists of $\mathcal{C} := \{(\mu_1, 0, 0), (\mu_2, 0, 0)\}$. Consider the following set of inefficient outcomes.

$$\mathcal{I}^+_{\mu_3} := \{(\mu_3, x, y) : x > 0, y > 0, x + y \leq 1\}$$

We argue that $\mathcal{C} \cup \mathcal{I}^+_{\mu_3}$ is a vNM stable set. By Observation 2, outcomes in $\mathcal{C}$ do not dominate each other. By Observation 1, outcomes in $\mathcal{I}^+_{\mu_3}$ do not dominate each other. Since firm f gets payoff 2 in outcomes in $\mathcal{C}$ and gets a maximum payoff of 1 in outcomes in $\mathcal{I}^+_{\mu_3}$, no outcome in $\mathcal{C}$ can be dominated by outcomes in $\mathcal{I}^+_{\mu_3}$. Finally, both workers get zero payoff in outcomes

in $\mathcal{C}$ but get positive payoff in outcomes in $\mathcal{I}^+_{\mu_3}$. Hence, outcomes in $\mathcal{I}^+_{\mu_3}$ cannot be dominated by outcomes in $\mathcal{C}$.

Next, we argue external stability. Outcomes corresponding to μ_4 are clearly dominated by $(\mu_1, 0, 0)$ (or any other outcome in $\mathcal{C} \cup \mathcal{I}^+_{\mu_3}$). Outcomes of the form (μ_3, x, y), where $xy = 0$, is dominated by $(\mu_1, 0, 0)$ if $x = 0$ and by $(\mu_2, 0, 0)$ if $y = 0$. ■

Appendix C contains another example in which the corresponding economy has an empty core but a vNM stable set containing *only inefficient* outcomes exist.

5 Toward a characterization of vNM stable sets

Example 1 shows that vNM stable sets may not exist with multiple efficient matchings. The following example shows that if there are multiple efficient matchings, $\mathcal{I}_\mu$ need not be vNM stable for any efficient matching μ. However, a vNM stable set exists in this example.

Example 3 (Multiple efficient matchings)**.** Consider the following economy. There is one firm f and two workers w and w'. The surplus function of the firm is $V_f(\{w\}) = V_f(\{w'\}) = 1$ and $V_f(\{w, w'\}) = 0$. There are two efficient matchings μ : matches the firm to worker w and μ' : matches the firm to worker w'. Then, it is not difficult to see that

$$\mathcal{I}_\mu = \{(\mu, u(w), u(w')) : u(w') = 0, 0 \leq u(w) \leq 1\}$$
$$\mathcal{I}_{\mu'} = \{(\mu', u(w), u(w')) : u(w) = 0, 0 \leq u(w') \leq 1\}$$

We first argue that $\mathcal{I}_\mu$ is not a vNM stable set. Indeed, the outcome $(\mu', 0, 0)$ cannot be dominated by any outcome in $\mathcal{I}_\mu$. To see this, if $(\mu, x, 0)$ dominates $(\mu', 0, 0)$, then we see that the payoff of firm is $1 - x \leq 1$, where 1 is its payoff in $(\mu', 0, 0)$. Hence, for domination, we must have $x = 0$ which gives equal payoff to firm f in both outcomes. Hence, all the workers are also indifferent between the two outcomes. Hence, $\mathcal{I}_\mu$ is not a vNM stable set. An analogous argument shows $\mathcal{I}_{\mu'}$ is also not a vNM stable set.

Is $\mathcal{I}_\mu \cup \mathcal{I}_{\mu'}$ a vNM stable set? No. Consider $(\mu, 0, 0)$ and $(\mu', 0, \frac{1}{2})$. Worker w gets the same payoff in both outcomes and firm f is strictly better in $(\mu, 0, 0)$. Hence, $(\mu, 0, 0)$ dominates $(\mu', 0, \frac{1}{2})$, violating internal stability.

Is anything vNM stable? Consider the set of two outcomes $\{(\mu, 0, 0), (\mu', 0, 0)\}$. The firm and the workers are all indifferent between the outcomes. So, they are not comparable, and

internally stable. For external stability, any other matching μ'' has only zero payoffs for firm and two workers.[9] Now, outcome $(\mu, 0, 0)$ dominates any such outcome. Also, $(\mu', 0, 0)$ dominates any outcome $(\mu, x, 0)$ with $0 < x \leq 1$. Similarly, $(\mu, 0, 0)$ dominates any outcome $(\mu', 0, x)$ with $0 < x \leq 1$. Hence, $\{(\mu, 0, 0), (\mu', 0, 0)\}$ is a vNM stable set. ■

We now impose a richness on vNM stable sets. We want our vNM stable sets to be rich in the sense that if an outcome (μ, u) is part of a vNM stable set then certain other outcomes corresponding to μ should be part of the vNM stable set. One reason why we investigate rich vNM stable sets is that it allows us to establish a partial converse of Theorem 1, and thus showing the salience of the rich vNM stable set identified in Theorem 1. The other reason to investigate rich vNM stable sets is methodological. As we show below, *rich* vNM stable sets are easier to analyze. Indeed, the sufficient condition (unique efficient matching) in Theorem 1 is almost necessary for the existence of rich vNM stable sets, and the rich vNM stable set identified in Theorem 1 is almost unique.

We introduce two version of richness, the first one weaker than the second. The first version of richness requires that certain firm-optimal outcomes to be included in the vNM stable sets.

Definition 3. *A set of outcomes $\mathcal{V} \subseteq \mathcal{I}$ is* **firm-rich** *if*

$$\Big[(\mu, u) \in \mathcal{V}\Big] \Longrightarrow \Big[(\mu, \mathbf{0}) \in \mathcal{V}\Big]$$

Notice that the outcome $(\mu, \mathbf{0})$ assigns zero payoff to all the workers and gives the entire surplus to firms. Hence, it corresponds to the *firm-optimal* outcome corresponding to matching μ. Firm-richness requires that if some IR outcome (μ, u) is included in $\mathcal{V}$, then the firm-optimal outcome corresponding to μ is also included in $\mathcal{V}$.

An economy is **generic** if for every $f \in F$ and every $S, T \subseteq W$, we have

$$\Big[V_f(S) = V_f(T)\Big] \Longrightarrow \Big[S = T\Big]$$

Note that the set of all efficient and IR outcomes is firm-rich. Our first characterization is of firm-rich vNM stable sets in generic economies.

[9] There are two other matchings, a matching that matches both the workers to the firm and a matching that matches no workers to the firm. Both these matchings generate zero surplus, and hence, by IR, must give zero payoff to the firm and workers.

Theorem 2. *Suppose the economy is generic. A set of outcomes $\mathcal{V}$ is a firm-rich vNM stable set if and only if*

(a) there is a unique efficient matching and

(b) $\mathcal{V}$ is the set of all efficient and IR outcomes.

Theorem 1 already proves one direction of Theorem 2. Hence, what remains to be proved is that existence of a firm-rich vNM stable set implies the economy must have a unique efficient matching, and that firm-rich vNM stable set is precisely the one identified in Theorem 1. We give a combined proof of Theorem 2 and 3 in Appendix A. To see how generic economy with firm-rich vNM stable set helps in the proof, let $\mathcal{V}$ be a firm-rich vNM stable set. Suppose $\mathcal{V}$ contains outcomes (μ, u) and (μ', u'). Firm-richness implies $(\mu, \mathbf{0}), (\mu', \mathbf{0}) \in \mathcal{V}$. Now, pick any firm f and suppose $\mu(f) \neq \mu'(f)$. If $V_f(\mu(f)) > V_f(\mu'(f))$, then $(\mu, \mathbf{0}) \xrightarrow{f} (\mu', \mathbf{0})$, which will contradict internal stability. Similarly, if $V_f(\mu(f)) < V_f(\mu'(f))$, then $(\mu', \mathbf{0}) \xrightarrow{f} (\mu, \mathbf{0})$. Hence, $V_f(\mu(f)) = V_f(\mu'(f))$ for all f. Generic economy assumption allows us to then conclude $\mu(f) = \mu'(f)$ for all f. Hence, we can conclude that in any firm-rich vNM stable set in a generic economy, outcomes corresponding to a single matching μ must exist. By Observation 1, such a vNM stable set must be $\mathcal{I}_\mu$ – the set of all IR outcomes corresponding to the matching μ. The fact that μ is the unique efficient matching is significantly more complicated to show. The proof of this part relies on an elegant result in network flows by David Gale, for which we interpret some of our external stability conditions existence of flows in an appropriate network.

Remark 2. Firm-richness requires inclusion of extreme outcomes in the vNM stable set. As a result, we show the almost necessity of unique efficient outcomes in vNM stable sets and an almost unique characterization of the vNM stable set identified in Theorem 1. However, it is arguable that such extreme outcomes should *not* be included in a vNM stable set, given that it is supposed to reflect "standard social behavior" (in the language of von Neumann and Morgenstern (1944)). While it is difficult to characterize all vNM stable sets, we believe that firm-richness is plausible. In economies where firms have high bargaining power, firm-optimal outcomes may have to be included in any menu of contracts. ■

We now relax the generic economy assumption. By requiring more richness in the vNM stable set, we derive a characterization analogous to Theorem 2.

We say a matching μ is **IR-feasible** if $\mathcal{I}_\mu$ is non-empty. For any set of outcomes $\mathcal{V} \subset \mathcal{I}$, denote by $\mathcal{M}(\mathcal{V})$ the set of matchings involved in $\mathcal{V}$:

$$\mathcal{M}(\mathcal{V}) := \{\mu : (\mu, u) \in \mathcal{V}\}$$

Without genericity, a firm-rich vNM stable set $\mathcal{V}$ may contain outcomes from more than one matching. To characterize such sets, we strengthen our richness requirement. We now consider worker-IR-maximal outcomes, and the consequence of including them in vNM stable sets. For every worker w, let

$$u(w) := \begin{cases} V_{\mu(w)}(\mu(\mu(w))) & \text{if } \mu(w) \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

and $u(w') := 0$ for all $w' \neq w$. Clearly, $(\mu, u) \in \mathcal{I}_\mu$. We refer to (μ, u) as w-**IR-maximal** outcome for μ.[10]

Definition 4. *A set of outcomes $\mathcal{V} \subset \mathcal{I}$ is* **rich** *if it is firm-rich and for every $\mu \in \mathcal{M}(\mathcal{V})$ and for every worker w, the w-IR-maximal outcome for μ belongs to $\mathcal{V}$.*

As we will show below, without genericity, a rich vNM stable set $\mathcal{V}$ may contain outcomes corresponding to more than one matching, but these matchings are essentially unique in the following sense.

Definition 5. *A set of IR-feasible matchings $\mathcal{M}'$ is* **essentially unique** *if for all $\mu, \mu' \in \mathcal{M}'$ and for all $f \in F$,*

$$\Big[\mu(f) \neq \mu'(f)\Big] \Longrightarrow \Big[V_f(\mu(f)) = V_f(\mu'(f)) = 0\Big]$$

For any matching μ, let $V(\mu) := \sum_{f \in F} V_f(\mu(f))$ denote the total surplus generated from μ. Note that if $\mathcal{M}'$ is essentially unique, then for all $\mu, \mu' \in \mathcal{M}'$, we have $V(\mu) = V(\mu')$.

Theorem 3. *A set of outcomes $\mathcal{V}$ is a rich vNM stable set if and only if*

(a) $\mathcal{M}(\mathcal{V})$ is essentially unique and

(b) $\mathcal{V}$ is the set of all efficient and IR outcomes.

[10] Notice that this outcome not only gives the maximum possible payoff to w but also gives zero payoff to all other workers. In that sense, it is choosing a w-IR-maximal outcome that is worst for other workers.

Thus, if we move beyond generic economies, we can establish the analogue of Theorem 2 under additional richness. The following is an immediate corollary of Theorem 3.

Corollary 2. *Suppose μ is an IR-feasible matching. Then, $\mathcal{V} \subseteq \mathcal{I}_\mu$ is vNM stable if and only if μ is the unique efficient matching and $\mathcal{V} = \mathcal{I}_\mu$.*

The proof of necessity direction of Corollary 2 follows because Observation 1 and external stability imply $\mathcal{V} = \mathcal{I}_\mu$, and hence, $\mathcal{V}$ is a rich vNM stable set. Theorem 3 then implies the result.

6 Conclusion

Our main motivation for studying vNM stable sets was that stable outcomes are not guaranteed to exist in this model, specially in a large class of economically meaningful settings where teams of workers exhibit strong complementarities. By extending this to a collective stability notion, we have *almost* resolved the existence issue. Our results suggest that in this model, if firms and workers were to choose a collection of outcomes (contracts) using the principles of vNM stability, they are *almost* forced to choose only efficient outcomes but have complete flexibility in the division of surplus.

References

Aumann, R. J. (1985): "What Is Game Theory Trying to Accomplish?" in *Frontiers of Economics*, ed. by K. J. Arrow and S. Honkapohja, Oxford: Basil Blackwell, 28–76.

Baldwin, E. and P. Klemperer (2019): "Understanding Preferences: 'Demand Types', and the Existence of Equilibrium With Indivisibilities," *Econometrica*, 87, 867–932.

Border, K. C. (1991): "Implementation of reduced form auctions: A geometric approach," *Econometrica*, 59, 1175–1187.

Che, Y.-K., J. Kim, and K. Mierendorff (2013): "Generalized Reduced-Form Auctions: A Network-Flow Approach," *Econometrica*, 81, 2487–2520.

Che, Y.-K., J. Kim, and F. Kojima (2013): "Stable matching in large economies," *Econometrica*, 87, 65–110.

Doğan, B. and L. Ehlers (2026): "Existence of Myopic-Farsighted Stable Sets in Matching Markets," *Econometrica*, 94, 811–841.

Ehlers, L. (2007): "Von Neumann–Morgenstern Stable Sets in Matching Problems," *Journal of Economic Theory*, 134, 537–547.

Hatfield, J. W., S. D. Kominers, A. Nichifor, M. Ostrovsky, and A. Westkamp (2019): "Full substitutability," *Theoretical Economics*, 14, 1535–1590.

Herings, P. J.-J., A. Mauleon, and V. Vannetelbosch (2017): "Stable Sets in Matching Problems with Coalitional Sovereignty and Path Dominance," *Journal of Mathematical Economics*, 71, 14–19.

——— (2020): "Matching with Myopic and Farsighted Players," *Journal of Economic Theory*, 190, 105125.

Kelso, A. S. and V. P. Crawford (1982): "Job matching, coalition formation, and gross substitutes," *Econometrica*, 1483–1504.

Klaus, B., F. Klijn, and M. Walzl (2011): "Farsighted Stability for Roommate Markets," *Journal of Public Economic Theory*, 13, 921–933.

Kojima, F., N. Sun, and N. N. Yu (2020): "Job matching under constraints," *American Economic Review*, 110, 2935–2947.

Liu, Q. (2024): "Matching with Incomplete Information," in *Handbook of the Economics of Matching*, vol. 1.

Liu, Q., G. J. Mailath, A. Postlewaite, and L. Samuelson (2014): "Stable Matching with Incomplete Information," *Econometrica*, 82, 541–587.
Rostek, M., and N. Yoder (2020): "Matching With Complementary Contracts," Econometrica, 88(5), 1793–1827. https://doi.org/10.3982/ECTA16686.

Rostek, M., and N. Yorder (2020): "Matching With Complementary Contracts," *Econometrica*, 88, 1793–1827.

Mauleon, A., V. J. Vannetelbosch, and W. Vergote (2011): "Von Neumann–Morgenstern Farsightedly Stable Sets in Two-Sided Matching," *Theoretical Economics*, 6, 499–521.

Núñez, M. and C. Rafels (2013): "Von Neumann–Morgenstern Solutions in the Assignment Market," *Journal of Economic Theory*, 148, 1282–1291.

Peters, H. (2008): *Game theory: A Multi-leveled approach*, Springer.

Pourpouneh, M., R. Ramezanian, A. Sen, and V. Taori (2026): "Stable Matching with Payoff Guarantees," ArXiv preprint.

Shapley, L. and M. Shubik (1971): "Assignment Game I: The Core," *International Journal of Game Theory*, 1, 111–130.

Subiza, B., J.-M. Giménez-Gómez, and J. E. Peris (2025): "Cooperative TU-games: Dominance, stable sets, and the core revisited," *Journal of Mathematical Economics*, 119, 103137.

von Neumann, J. and O. Morgenstern (1944): *Theory of Games and Economic Behavior*, Princeton University Press.

Wako, J. (2010): "A Polynomial-Time Algorithm to Find von Neumann–Morgenstern Stable Matchings in Marriage Games," *Algorithmica*, 58, 188–220.

A Proofs of Theorems 2 and 3

We give a proof of Theorems 2 and 3 here. The main ingredient of the proofs is the following proposition about the domination relation, which does not assume any kind of richness, and is of indepenendent interest.

Proposition 1. *Suppose $\mu^\star$ is an efficient matching and μ is an IR-feasible matching. Then, there exists $(\mu^\star, u^\star) \in \mathcal{I}_{\mu^\star}$ such that*

$$(\mu, u) \nrightarrow (\mu^\star, u^\star) \qquad \forall\ (\mu, u) \in \mathcal{I}_\mu$$

The proof of Proposition 1, which is given below after some preparations for the proof is done, hinges on a characterization of the conditions under which an outcome $(\mu', u') \in \mathcal{I}$ is not dominated by any outcome $(\mu, u) \in \mathcal{I}_\mu$, where $\mu \neq \mu'$. This is done in Lemma 3, whose proof is in Appendix B.

Lemma 3. *Let μ' and μ be two distinct IR-feasible matchings. Suppose $(\mu', u') \in \mathcal{I}_{\mu'}$. Then, the following are equivalent:*

1. *$(\mu, u) \nrightarrow (\mu', u')$ for all $(\mu, u) \in \mathcal{I}_\mu$*
2. *For all $f \in F$,*

$$V_f(\mu(f)) - \sum_{w \in \mu(f)} u'(w) \leq V_f(\mu'(f)) - \sum_{w \in \mu'(f)} u'(w) \tag{1}$$

Fix any matching μ'. Lemma 3 gives a clean linear system to verify when there exists u' such that $(\mu', u') \in \mathcal{I}_{\mu'}$ and $(\mu, u) \nrightarrow (\mu', u')$ for every $(\mu, u) \in \mathcal{I}_\mu$. We rewrite this linear system below.

To do so, we introduce a dummy firm f_0 such that $V_{f_0}(S) = 0$ for all $S \subset W$. We assume that any worker who is unmatched in a matching is matched to f_0. Without loss of generality, we assume that $f_0 \in F$ and each worker is matched to some firm in F (possibly to f_0) in every matching, where $\mu(w) = f_0$ for a matching μ would indicate that worker w is unmatched. Now, consider the following system of inequalities.

$$\sum_{w \in \mu'(f)} u'(w) - \sum_{w \in \mu(f)} u'(w) \leq V_f(\mu'(f)) - V_f(\mu(f)) \qquad \forall\ f \in F \tag{\textbf{ES}}$$

$$\sum_{w \in \mu'(f)} u'(w) \leq V_f(\mu'(f)) \qquad \forall\ f \in F \tag{IR}$$
$$u'(w) \geq 0 \qquad \forall\ w \in W$$

Constraints (**IR**) ensures that a feasible u' satisfies $(\mu', u') \in \mathcal{I}$. Constraints (**ES**) is the same as the constraints in Lemma 3. Note that if a worker w is unmatched in μ' (that is matched to f_0), then $u'(w) \geq 0$ and (**IR**) along with the fact that $V_{f_0}(\mu'(f_0)) = 0$ implies that $u'(w) = 0$.

If u' is a feasible solution to (**ES**) and (**IR**), then there exists $(\mu', u') \in \mathcal{I}_{\mu'}$ such that $(\mu, u) \nrightarrow (\mu', u')$ for all $(\mu, u) \in \mathcal{I}_\mu$. Conversely, if there exists $(\mu', u') \in \mathcal{I}_{\mu'}$, it is already a feasible solution to (**IR**) and if it is such that $(\mu, u) \nrightarrow (\mu', u')$ for all $(\mu, u) \in \mathcal{I}_\mu$, then by Lemma 3, it is a feasible solution to (**ES**).

To verify feasibility of (**ES**) and (**IR**) we represent this as a *demand graph*, whose flows will represent a feasible solution. Our demand graph $\mathcal{G}$ has the following components: $\mathcal{G} \equiv (V, E, c, b)$ is a **demand graph**, where

- V is a finite set of vertices
- $E \subseteq \{(x, y) : x, y \in V\}$ is a finite set of directed edges
- $c : E \to \Re_+$ specifies a non-negative capacity for each edge in E
- $b : V \to \Re$ specifies *divergence* at each vertex – if $b(x) > 0$, then x is a *demand* vertex and if $b(x) < 0$, then x is a *supply* vertex

A *flow* in $\mathcal{G}$ is a map $\phi : E \to \Re_+$ satisfying the following feasibility constraints:

$$\phi(x, y) \leq c(x, y) \qquad \forall\ (x, y) \in E$$
$$\underbrace{\sum_{(y,x) \in E} \phi(y, x)}_{\text{In flow}} - \underbrace{\sum_{(x,y) \in E} \phi(x, y)}_{\text{Out flow}} = b(x) \qquad \forall\ x \in V$$

We fix matchings μ and μ' and construct the specific demand graph $\mathcal{G}$ as follows:

- **Vertices.** There is a vertex for each worker $w \in W$. Denote such a vertex as w itself for each $w \in W$. There is a pair of vertices for each firm $f \in F$. Denote such vertices as $f, \hat{f}$ for each firm $f \in F$. The set of vertices $\hat{f}$ for each f will be referred to as

the auxiliary vertices corresponding to firms. We will denote the set of all auxiliary vertices as $\widehat{F}$. There is another vertex v_0 for handling slacks in constraints (**ES**). So, in total there are $|W| + 2|F| + 1$ vertices.

- **Edges.** For every $w \in W$, there is an edge from w to f if $\mu'(w) = f$. There is an edge from f to $\hat{f}$ for each $f \in F$. For each $f \in F$, there is an edge from $\hat{f}$ to w if $\mu(w) = f$. There is an edge from v_0 to $\hat{f}$ for every $f \in F$. Note that each vertex w has exactly one outgoing edge (to $\mu'(w)$) and exactly one incoming edge (from $\mu(w)$).[11]

- **Capacities.** All edges have very high capacity (to be precise more than $V(\mu')$) except edges $(f, \hat{f})$ for each $f \in F$:

$$c(f, \hat{f}) := V_f(\mu'(f))$$

- **Divergences.** Divergence at vertex $w \in W$ is zero: $b(w) = 0$. Divergence at the two vertices corresponding to any firm $f \in F$ are: $b(f) = 0$ but

$$b(\hat{f}) := V_f(\mu'(f)) - V_f(\mu(f))$$

  Divergence at v_0 is

$$b(v_0) := V(\mu) - V(\mu')$$

  where for any matching $\hat{\mu}$, $V(\hat{\mu}) \equiv \sum_{f \in F} V_f(\hat{\mu}(f))$.

  Note that $\sum_{x \in V} b(x) = 0$.

As an example consider $W = \{w_1, w_2, w_3, w_4\}$ and $F = \{f_1, f_2, f_3\}$. Suppose μ and μ' are as follows:

$$\mu'(w_1) = f_1, \mu'(w_2) = f_1, \mu'(w_3) = f_2, \mu'(w_4) = f_3$$
$$\mu(w_1) = f_3, \mu(w_2) = f_2, \mu(w_3) = f_1, \mu(w_4) = f_2$$

The corresponding demand graph is shown in Figure 1

The following lemma makes the connection between demand graph and the system of inequalities in (**ES**) and (**IR**) precise.

[11] This is because each worker is matched to some firm, possible f_0.

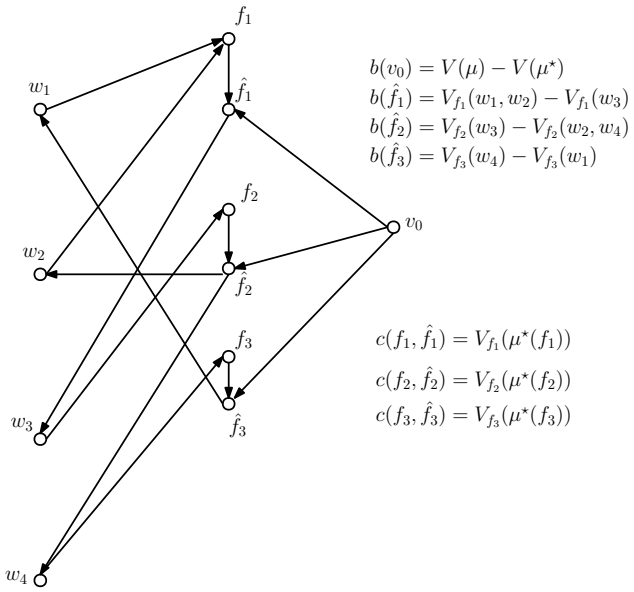


Figure 1: Demand network

Lemma 4. *Suppose μ and μ' are two IR-feasible matchings. A feasible solution to (**ES**) and (**IR**) exists if and only if a flow in the demand graph $\mathcal{G}$ exists.*

We use Lemma 4, which is proved in Appendix B, to show sufficiency of efficient matching for (**ES**) and (**IR**) to be feasible. The proof of Proposition 1 can now be completed using a celebrated theorem of David Gale on existence of flows in a demand graph.[12]

Proof of Proposition 1. Let μ and $\mu^\star \equiv \mu'$ be two IR-feasible matchings (to be consistent with the notation in the earlier lemmas, we will denote $\mu' \equiv \mu^\star$). We show that if μ' is an efficient matching, then a flow in directed graph $\mathcal{G}$ (corresponding to μ and $\mu^\star$) exists. Lemma 4 will then imply that a feasible solution to (**ES**) and (**IR**) exists. Denote this solution as $(\mu^\star, u^\star)$. and Lemma 3 will imply that for all $(\mu, u) \in \mathcal{I}_\mu$, we have $(\mu, u) \nrightarrow (\mu^\star, u^\star)$. This will complete the proof.

To show that a flow in directed graph $\mathcal{G}$ exists, we invoke a seminal result due to David Gale for this. We introduce some notations first. A *cut* in $\mathcal{G}$ is a partitioning of the set of vertices into two parts: $(S, V \setminus S)$. Without loss of generality, we denote a cut by $S \subseteq V$.

[12] David Gale's existence theorem has been used to prove an important extension of Border's reduced form auction result (Border, 1991) – see Che et al. (2013).

The capacity of a cut is the sum of capacities of all incoming edges to the cut:

$$c(S) := \sum_{(x,y)\in E: x\notin S, y\in S} c(x,y) \qquad \forall\ S \subseteq V$$

The divergence of a cut is the sum of divergences of vertices in the cut:

$$b(S) := \sum_{x\in S} b(x) \qquad \forall\ S \subseteq V$$

Theorem 4 (David Gale's Demand Theorem)**.** *Suppose $\mathcal{G}$ is a directed graph such that $b(V) = 0$. Then, a flow exists in $\mathcal{G}$ if and only if for every cut $S \subseteq V$, we have*

$$b(S) \leq c(S)$$

By David Gale's theorem, it is enough to show that for any cut $S \subseteq V$ in $\mathcal{G}$, we have $b(S) \leq c(S)$. We will consider three kinds of cuts.

Case 1. This is the case where there is some incoming edge to S with high capacity. Hence, $c(S)$ is a very high number and $c(S) \geq b(S)$.

Case 2. There is no high capacity edge incoming to S and $v_0 \notin S$. In that case, each auxiliary vertex $\hat{f} \notin S$ – else there would be a high capacity edge $(v_0, \hat{f})$ into S. Since divergence is non-zero only for vertices in $\widehat{F} \cup \{v_0\}$, we conclude that $b(S) = 0$. Since capacities are non-negative, $c(S) \geq b(S)$.

Case 3. There is no high capacity edge incoming to S and $v_0 \in S$. We are going to partition auxiliary (firm)-vertices in S into two parts: $(\widehat{F}' \cup \widehat{F}'') \subset S$ with $\widehat{F}' \cap \widehat{F}'' = \emptyset$. For every $\hat{f} \in \widehat{F}'$, the vertex $f \in S$, and hence, edge $(f, \hat{f})$ is not incoming to S. For every $\hat{f} \in \widehat{F}''$, vertex $f \notin S$, and hence, edge $(f, \hat{f})$ is incoming to S.[13] That is $\widehat{F}'' \subset S$ and $F'' \subset V \setminus S$. Partition the set of workers also:

$$W' := \bigcup_{f'\in F'} \mu'(f'), \qquad W'' := \bigcup_{f''\in F''} \mu'(f''), \qquad W \setminus (W' \cup W'')$$

[13]This case also covers the case where there are no incoming edges to S, and hence, $c(S) = 0$. In that case $F'' = \emptyset$.

Since S has no incoming-edge with high capacity, two things must hold true:

- $\mu'(w) \in F \setminus (F' \cup F'')$ for all $w \in W \setminus (W' \cup W'')$,
- $W' \subset S$ and for every $w \in W'$, $\mu(w) \in F' \cup F''$.

An illustration of cut S is shown in Figure 2.

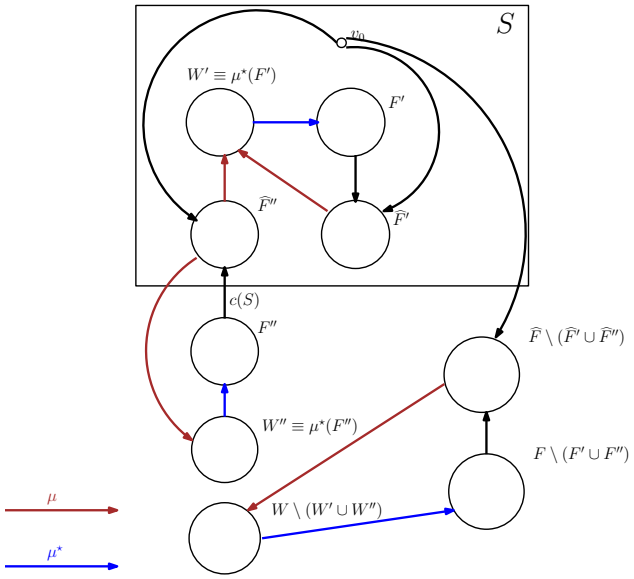


Figure 2: Representation of cut S

The divergence of S equals

$$\begin{aligned} b(S) &= b(v_0) + \sum_{\hat{f} \in \widehat{F}'} b(\hat{f}) + \sum_{\hat{f} \in \widehat{F}''} b(\hat{f}) \\ &= b(v_0) + \sum_{f \in F'} \big[V_f(\mu'(f)) - V_f(\mu(f))\big] + \sum_{f \in F''} \big[V_f(\mu'(f)) - V_f(\mu(f))\big] \\ &= \big[V(\mu) - V(\mu')\big] + \sum_{f \in F'} \big[V_f(\mu'(f)) - V_f(\mu(f))\big] + \sum_{f \in F''} \big[V_f(\mu'(f)) - V_f(\mu(f))\big] \\ &= \sum_{f \in F \setminus (F' \cup F'')} \big[V_f(\mu(f)) - V_f(\mu'(f))\big]. \end{aligned}$$

The capacity of S equals $c(S) = \sum_{f \in F''} c(f, \hat{f}) = \sum_{f \in F''} V_f(\mu'(f))$ as the only incoming edges to S are of the form $(f, \hat{f})$ where $f \in F''$.

Consider a candidate matching $\hat{\mu}$ defined as follows. For every $f \in F \setminus \{f_0\}$, let

$$\hat{\mu}(f) = \begin{cases} \mu'(f) & \text{if } f \in F', \\ \emptyset & \text{if } f \in F'', \\ \mu(f) & \text{otherwise.} \end{cases}$$

All remaining workers are matched to the dummy firm f_0, that is, $\hat{\mu}(f_0) = W \setminus \big(\bigcup_{f \in F \setminus \{f_0\}} \hat{\mu}(f) \big)$.

Since for every $w \in W'$, $\mu(w) \in F' \cup F''$, the matching $\hat{\mu}$ is well-defined. As μ' is efficient,

$$0 \geq V(\hat{\mu}) - V(\mu') = \sum_{f \in F \setminus (F' \cup F'')} \big[V_f(\mu(f)) - V_f(\mu'(f)) \big] - \sum_{f \in F''} V_f(\mu'(f)) = b(S) - c(S).$$

Hence, it follows that $b(S) \leq c(S)$. ■

We now have all the tools to prove Theorems 2 and 3.

Proof of Theorems 2 and 3.

Sufficiency: The sufficiency direction of Theorem 3 implies the sufficiency direction of Theorem 2. Hence, we only give a proof of sufficiency direction of Theorem 3. Let $\mathcal{M}^{\text{eff}}$ be the set of all efficient matchings. Suppose $\mathcal{M}^{\text{eff}}$ is essentially unique, and let

$$\mathcal{V} := \bigcup_{\mu \in \mathcal{M}^{\text{eff}}} \mathcal{I}_\mu$$

We show $\mathcal{V}$ is a vNM stable set in two steps.

Internal stability. Consider any two efficient matchings $\mu, \mu' \in \mathcal{M}^{\text{eff}}$. Consider any $f \in F$. Since $\mathcal{M}^{\text{eff}}$ is essentially unique, $V_f(\mu(f)) = V_f(\mu'(f))$. If $\mu(f) = \mu'(f)$, then for every $(\mu', u') \in \mathcal{I}_{\mu'}$, we have

$$V_f(\mu(f)) - \sum_{w \in \mu(f)} u'(w) = V_f(\mu'(f)) - \sum_{w \in \mu'(f)} u'(w) \tag{2}$$

Lemma 3 implies $(\mu, u) \nrightarrow (\mu', u')$ for all $(\mu, u) \in \mathcal{I}_\mu$. If $\mu(f) \neq \mu'(f)$, then $V_f(\mu(f)) = V_f(\mu'(f)) = 0$ (by essential uniqueness). Hence, for every $(\mu', u') \in \mathcal{I}_{\mu'}$, we have $u'(w) = 0$ for all $w \in \mu'(f)$ and using the fact that $u'(w) \geq 0$ for all w, equation (2) holds with in-

equality (LHS $\leq$ RHS). Hence, Lemma 3 implies $(\mu, u) \nrightarrow (\mu', u')$ for all $(\mu, u) \in \mathcal{I}_\mu$. Finally, if $\mu = \mu'$, Observation 1 implies $(\mu, u) \nrightarrow (\mu', u')$ for all $(\mu, u) \in \mathcal{I}_\mu$. This shows that $\mathcal{V}$ is internally stable.

External stability. Consider any outcome $(\mu, u) \in \mathcal{I} \setminus \mathcal{V}$. Hence, $\mu \notin \mathcal{M}^{\text{eff}}$. By Lemma 2, there exists an efficient matching $\mu^\star \in \mathcal{M}^{\text{eff}}$ and an outcome $(\mu^\star, u^\star) \in \mathcal{I}_{\mu^\star}$ such that $(\mu^\star, u^\star) \to (\mu, u)$. Hence, external stability follows.

Necessity: We prove the necessity directions of Theorems 2 and 3 together. Suppose $\mathcal{V}$ is a vNM-stable set. The following claim clarifies the role firm-richness in a vNM stable set.

Claim 1. *Suppose $\mathcal{V}$ is a firm-rich vNM-stable set. Then, for every $\mu, \mu' \in \mathcal{M}(\mathcal{V})$, we have*

$$V_f(\mu(f)) = V_f(\mu'(f)) \qquad \forall\ f \in F \tag{3}$$

Further, if the economy is generic $\mu = \mu'$ (i.e., $\mathcal{M}(\mathcal{V})$ has a unique matching).

Proof. Suppose $\mathcal{V}$ is a firm-rich vNM-stable set. Consider two distinct matchings $\mu, \mu' \in \mathcal{M}(\mathcal{V})$. Since $\mathcal{I}_\mu$ and $\mathcal{I}_{\mu'}$ are non-empty, individual rationality implies $V_f(\mu(f)) \geq 0$ and $V_f(\mu'(f)) \geq 0$ for every $f \in F$. Consider u and u' such that $u(w) = u'(w) = 0$ for every $w \in W$. Note that u is firm-IR-optimal for μ and u' is firm-IR-optimal for μ'. Hence, we have $(\mu, u), (\mu', u') \in \mathcal{V}$. If for some $f \in F$, $V_f(\mu(f)) > V_f(\mu'(f))$, then $(\mu, u) \xrightarrow{f} (\mu', u')$, contradicting internal stability. Hence, $V_f(\mu(f)) \leq V_f(\mu'(f))$ for all $f \in F$. By interchanging the roles of (μ, u) and (μ', u'), we obtain $V_f(\mu'(f)) \leq V_f(\mu(f))$ for all $f \in F$. Therefore, (3) holds.

If the economy is generic, by definition for every f, we have $\mu(f) = \mu'(f)$, and hence, $\mu = \mu'$. □

By Claim 1, if $\mathcal{V}$ is firm-rich, then for every $\mu, \mu' \in \mathcal{M}(\mathcal{V})$, we have

$$V_f(\mu(f)) = V_f(\mu'(f)) \qquad \forall\ f \in F \tag{4}$$

Further, if the economy is generic, then $\mathcal{M}(\mathcal{V})$ contains a unique matching.

If the economy is not generic, then assume $\mathcal{V}$ is rich. Suppose now that for some $f \in F$ we

have $V_f(\mu(f)) = V_f(\mu'(f)) = p > 0$ but $\mu(f) \neq \mu'(f)$.[14] Without loss of generality, suppose there is some $\tilde{w} \in \mu(f) \setminus \mu'(f)$. Let $\tilde{u}$ be a utility vector such that $\tilde{u}(\tilde{w}) = p, \tilde{u}(w) = 0$ for all $w \neq \tilde{w}$. Clearly, $\tilde{u}$ is such that $(\mu, \tilde{u}) \in \mathcal{I}_\mu$. Further, $(\mu, \tilde{u})$ is $\tilde{w}$-IR-maximal for μ. Hence, by richness, $(\mu, \tilde{u}) \in \mathcal{V}$. Moreover, all workers in $\mu'(f)$ receive zero payoff in (μ', u'), while firm f receives payoff p at (μ', u') and payoff zero at $(\mu, \tilde{u})$. Hence, $(\mu', u') \xrightarrow{f} (\mu, \tilde{u})$, contradicting internal stability (since both $(\mu', u'), (\mu, \tilde{u}) \in \mathcal{V}$). Thus, $V_f(\mu(f)) = V_f(\mu'(f)) > 0$ implies that $\mu(f) = \mu'(f)$. Along with (4), this proves that $\mathcal{M}(\mathcal{V})$ is essentially unique.

We show that $\mathcal{M}(\mathcal{V}) = \mathcal{M}^{\text{eff}}$ (in case of generic economy, this will show that $\mathcal{M}(\mathcal{V})$ contains a unique efficient matching, as is necessary of Theorem 2). First, note that by adding (4) for all $f \in F$, we see that

$$V(\mu) = V(\mu') \qquad \forall\ \mu, \mu' \in \mathcal{M}(\mathcal{V}) \tag{5}$$

In other words, all matchings in $\mathcal{M}(\mathcal{V})$ generate the same surplus.

Suppose that there exists $\mu^\star \in \mathcal{M}^{\text{eff}} \setminus \mathcal{M}(\mathcal{V})$. Consider $\bar{\mu} \in \mathcal{M}(\mathcal{V})$. By Proposition 1, there exists $(\mu^\star, u^\star) \in \mathcal{I}_{\mu^\star}$ such that $(\bar{\mu}, \bar{u}) \nrightarrow (\mu^\star, u^\star)$ for all $(\bar{\mu}, \bar{u}) \in \mathcal{I}_{\bar{\mu}}$.

We now show that this implication spreads to all $\mu \in \mathcal{M}(\mathcal{V})$ since $\mathcal{M}(\mathcal{V})$ is essentially unique. To see this, consider any $\mu \in \mathcal{M}(\mathcal{V})$. Since $(\bar{\mu}, \bar{u}) \nrightarrow (\mu^\star, u^\star)$ for all $(\bar{\mu}, \bar{u}) \in \mathcal{I}_{\bar{\mu}}$, by Lemma 3, for all $f \in F$,

$$V_f(\bar{\mu}(f)) - \sum_{w \in \bar{\mu}(f)} u^\star(w) \leq V_f(\mu^\star(f)) - \sum_{w \in \mu^\star(f)} u^\star(w)$$

Consider any $f \in F$. If $\mu(f) = \bar{\mu}(f)$, then

$$V_f(\mu(f)) - \sum_{w \in \mu(f)} u^\star(w) = V_f(\bar{\mu}(f)) - \sum_{w \in \bar{\mu}(f)} u^\star(w) \leq V_f(\mu^\star(f)) - \sum_{w \in \mu^\star(f)} u^\star(w)$$

If $\mu(f) \neq \bar{\mu}(f)$, then essential uniqueness implies $V_f(\mu(f)) = V_f(\bar{\mu}(f)) = 0$. Hence,

$$V_f(\mu(f)) - \sum_{w \in \mu(f)} u^\star(w) = -\sum_{w \in \mu(f)} u^\star(w) \leq 0$$

[14] If our economy is without ties as in Remark 2, we do not need this part as $\mathcal{M}(\mathcal{V})$ will have only one matching, and we can establish essential uniqueness with weaker richness requirement: firm-IR-optimal outcomes are required to be included in $\mathcal{V}$.

whereas individual rationality of $(\mu^\star, u^\star)$ implies

$$V_f(\mu^\star(f)) - \sum_{w \in \mu^\star(f)} u^\star(w) \geq 0$$

Thus, in either case,

$$V_f(\mu(f)) - \sum_{w \in \mu(f)} u^\star(w) \leq V_f(\mu^\star(f)) - \sum_{w \in \mu^\star(f)} u^\star(w)$$

By Lemma 3, no outcome in $\mathcal{I}_\mu$ dominates $(\mu^\star, u^\star)$. Since this is true for every $\mu \in \mathcal{M}(\mathcal{V})$, no outcome in $\mathcal{V}$ dominates $(\mu^\star, u^\star)$:

$$(\mu, u) \nrightarrow (\mu^\star, u^\star) \qquad \forall\ (\mu, u) \in \mathcal{V}$$

Since $(\mu^\star, u^\star) \notin \mathcal{V}$, this contradicts external stability. Hence, every efficient matching belongs to $\mathcal{M}(\mathcal{V})$. Using (5), we conclude that $\mathcal{M}(\mathcal{V}) = \mathcal{M}^{\text{eff}}$.

Finally, we argue that $\mathcal{V} = \bigcup_{\mu \in \mathcal{M}^{\text{eff}}} \mathcal{I}_\mu$ (if the economy is generic, this will show that $\mathcal{V}$ contains the set of efficient and IR outcomes, as is necessary for Theorem 2). We have already established that $\mathcal{V} \subset \bigcup_{\mu \in \mathcal{M}^{\text{eff}}} \mathcal{I}_\mu$. Assume for contradiction that some efficient outcome $(\mu^\star, u^\star) \notin \mathcal{V}$. Then, by external stability, there is some efficient outcome $(\mu, u) \in \mathcal{V}$ and a firm f such that $(\mu, u) \xrightarrow{f} (\mu^\star, u^\star)$:

$$\begin{aligned} u(w) &\geq u^\star(w) \qquad \forall\ w \in \mu(f) \\ V_f(\mu(f)) - \sum_{w \in \mu(f)} u(w) &\geq V_f(\mu^\star(f)) - \sum_{w \in \mu^\star(f)} u^\star(w) \end{aligned}$$

with one inequality strict. Since $\mathcal{M}^{\text{eff}}$ is essentially unique, we get $V_f(\mu(f)) = V_f(\mu^\star(f))$. Using this and adding all the previous inequalities, we get

$$\sum_{w \in \mu^\star(f)} u^\star(w) > \sum_{w \in \mu(f)} u^\star(w)$$

If $\mu(f) = \mu^\star(f)$, this is a contradiction. Else, by essential uniqueness, we have $V_f(\mu(f)) = V_f(\mu^\star(f)) = 0$. But then, IR implies that $u^\star(w) = 0$ for all $w \in \mu(f)$ and for all $w \in \mu^\star(f)$. This will again contradict the previous inequality. ■

B Omitted Proofs of Lemmas

Proof of Lemma 1: Let $\mathcal{V} \subseteq \mathcal{I}$ be a candidate vNM stable set. We consider multiple steps.

Step 1. Consider a matching $\mu \notin \{\mu_1, \mu_2, \mu_3\}$. We know that if $\mu \notin \{\mu_1, \mu_2, \mu_3\}$ and $(\mu, u_{w_1}, u_{w_2}, u_{w_3}) \in \mathcal{I}$, then $u_{w_k} = 0$ for all $k \in \{1, 2, 3\}$. Hence, the only outcome corresponding to μ which may belong to $\mathcal{V}$ is $(\mu, 0, 0, 0)$. Consider $(\mu_1, x, y, 0) \in \mathcal{I}$ where $x + y \leq 1$ and $x, y \geq 0$. In this case, $(\mu_1, x, y, 0) \xrightarrow{f_1} (\mu, 0, 0, 0)$. A similar argument shows that $(\mu_2, 0, x, y) \xrightarrow{f_2} (\mu, 0, 0, 0)$ and $(\mu_3, x, 0, y) \xrightarrow{f_3} (\mu, 0, 0, 0)$. Hence, if $(\mu, 0, 0, 0) \in \mathcal{V}$ for some $\mu \notin \{\mu_1, \mu_2, \mu_3\}$, internal stability requires that no outcomes involving μ_1, μ_2, μ_3 can be in $\mathcal{V}$. Hence, external stability requires that some $(\mu, 0, 0, 0)$, where $\mu \notin \{\mu_1, \mu_2, \mu_3)$, must dominate $(\mu_1, x, y, 0)$, which is not possible. Thus, $\mathcal{V}$ only contains outcomes corresponding to efficient matchings μ_1, μ_2, μ_3.

Step 2. The outcome $(\mu_1, 1, 0, 0)$ cannot be dominated by $(\mu_3, x, 0, y)$. This is because if $(\mu_3, x, 0, y)$ were to dominate $(\mu_1, 1, 0, 0)$, then $x = 1$ and hence, $y = 0$, So, $(\mu_3, 1, 0, 0)$ needs to dominate $(\mu_1, 1, 0, 0)$, which is not possible because all firms get zero payoff in both the outcomes. Hence, it is not possible $\mathcal{V} = \mathcal{I}_{\mu_3}$. A similar argument extends this conclusion: it is not possible that $\mathcal{V} = \mathcal{I}_{\mu_i}$ for any $i \in \{1, 2, 3\}$. By Observation 1, $\mathcal{V} \nsubseteq \mathcal{I}_{\mu_i}$ for any $i \in \{1, 2, 3\}$.

Step 3. So, $\mathcal{V}$ must contain an outcome each from $\mathcal{I}_{\mu_i}$ for at least two values of i. Consider, without loss of generality, two outcomes $a_1 \equiv (\mu_1, x, y, 0) \in \mathcal{I}$ and $a_2 \equiv (\mu_2, 0, y', z') \in \mathcal{I}$. Note that if $y > y'$, then $a_1 \to a_2$ and if $y' > y$, then $a_2 \to a_1$. So, for internal stability, $y = y'$, i.e., $a_2 = (\mu_2, 0, y, z')$. But if $z' > 0$, we have $a_2 \to a_1$. Hence, internal stability forces $z' = 0$. Similarly, $x = 0$. So, $a_1 = (\mu_1, 0, y, 0)$ and $a_2 = (\mu_2, 0, y, 0)$. If $y < 1$, then firm 1 gets positive payoff in a_1 (and firm 2 gets positive payoff in a_2), and $a_1 \to a_2$ in that case. So, internal stability forces $y = 1$. Hence, for internal stability, the only outcomes in $\mathcal{I}_{\mu_1}$ and $\mathcal{I}_{\mu_2}$ that can be in $\mathcal{V}$ are $(\mu_1, 0, 1, 0)$ and $(\mu_2, 0, 1, 0)$.

Now, consider $a_3 \equiv (\mu_3, x'', 0, z'')$. Can $a_3 \in \mathcal{V}$? Since $a_1 \in \mathcal{V}$, we argue that $a_3 \notin \mathcal{V}$ due to internal stability. This is because $a_3 \to a_1$ (since firm 3, and workers 1 and 3 get zero payoff in a_1, and at least one of them must get positive payoff in a_3). Hence $a_3 \notin \mathcal{V}$ for any values of $1 \geq x'' \geq 0$ and $1 \geq z'' \geq 0$. So, the only kind of outcomes in $\mathcal{V}$ has one of the

matchings in $\{\mu_1, \mu_2, \mu_3\}$ and assigns the total surplus 1 to one of the workers matched.

But then, consider $\hat{a}_3 \equiv (\mu_3, \frac{1}{2}, 0, \frac{1}{2})$. For external stability, either a_1 or a_2 must dominate $\hat{a}_3$. But a_1 cannot dominate $\hat{a}_3$ because worker 1 gets positive payoff in $\hat{a}_3$ but gets zero in a_1. Similarly, a_2 cannot dominate $\hat{a}_3$ because worker 3 gets positive payoff in $\hat{a}_3$ but gets zero in a_2. By Observation 1, no outcome involving μ_3 can dominate $\hat{a}_3$. This implies that $\mathcal{V}$ fails to satisfy external stability. ■

Proof of Lemma 2: We do the proof in two steps.

STEP 1. We show that there exists some firm f such that $\mu(f) \neq \mu^\star(f)$ and

$$V_f(\mu^\star(f)) > V_f(\mu(f)) - \sum_{w \in \mu(f)} u(w) + \sum_{w \in \mu^\star(f)} u(w) \tag{6}$$

If this is not true, the reverse inequality holds for all firms whose matchings differ in μ and $\mu^\star$. Adding them together along with the firms whose matchings coincide (note the inequality (6) is an equality if $\mu(f) = \mu^\star(f)$) gives

$$\begin{aligned}
\sum_{f \in F} V_f(\mu^\star(f)) &\leq \sum_{f \in F} V_f(\mu(f)) - \sum_{f \in F} \sum_{w \in \mu(f)} u(w) + \sum_{f \in F} \sum_{w \in \mu^\star(f)} u(w) \\
&= \sum_{f \in F} V_f(\mu(f)) - \sum_{w \in W} u(w) + \sum_{f \in F} \sum_{w \in \mu^\star(f)} u(w) \\
&\leq \sum_{f \in F} V_f(\mu(f)) - \sum_{w \in W} u(w) + \sum_{w \in W} u(w) \\
&= \sum_{f \in F} V_f(\mu(f)),
\end{aligned}$$

where the first equality follows from the fact that unmatched workers in (μ, u) must get zero payoff (IR) and the second inequality follows because $u(w) \geq 0$ for all $w \in W$ (again IR). But this contradicts the fact that μ is not an efficient matching and $\mu^\star$ is efficient. Hence, there exists a firm f such that $\mu(f) \neq \mu^\star(f)$ and Inequality (6) holds. Let

$$\Delta := V_f(\mu^\star(f)) - V_f(\mu(f)) + \sum_{w \in \mu(f)} u(w) - \sum_{w \in \mu^\star(f)} u(w) \tag{7}$$

By Inequality (6), we have $\Delta > 0$.

STEP 2. We construct $u^\star$ such that $(\mu^\star, u^\star) \in \mathcal{I}_{\mu^\star}$ and $(\mu^\star, u^\star) \xrightarrow{f} (\mu, u)$, where f is the identified firm in Step 1. The $u^\star$ is defined as follows:

$$u^\star(w) = \begin{cases} u(w) + \frac{\Delta}{|\mu^\star(f)|} & \text{if } w \in \mu^\star(f) \\ 0 & \text{otherwise} \end{cases}$$

Since $u(w) \geq 0$, we have $u^\star(w) \geq 0$. Further, $u^\star(w) > u(w)$ for all $w \in \mu^\star(f)$. The utility of every firm $f' \neq f$ from the outcome $(\mu^\star, u^\star)$ is $V_{f'}(\mu^\star(f'))$ (since $u^\star(w) = 0$ for all $w \in \mu^\star(f')$), which is non-negative.[15] Finally, the utility of firm f from outcome $(\mu^\star, u^\star)$ is

$$V_f(\mu^\star(f)) - \sum_{w \in \mu^\star(f)} u^\star(w) = V_f(\mu^\star(f)) - \sum_{w \in \mu^\star(f)} u(w) - \Delta = V_f(\mu(f)) - \sum_{w \in \mu(f)} u(w) \tag{8}$$

where the last equality follows from (7). Hence, the payoff of firm f is the same in the two outcomes: (μ, u) and $(\mu^\star, u^\star)$. Since $(\mu, u) \in \mathcal{I}$, we thus conclude that $(\mu^\star, u^\star) \in \mathcal{I}$. Further, since each worker in $\mu^\star(f)$ is better off in $(\mu^\star, u^\star)$ than (μ, u), we conclude that $(\mu^\star, u^\star) \xrightarrow{f} (\mu, u)$. ■

Proof of Lemma 3. (2) $\Rightarrow$ (1). Suppose for some $(\mu, u) \in \mathcal{I}_\mu$, we have $(\mu, u) \xrightarrow{f} (\mu', u')$ for some firm f. Then,

$$\begin{aligned} u(w) &\geq u'(w) \qquad \forall\ w \in \mu(f) \\ V_f(\mu(f)) - \sum_{w \in \mu(f)} u(w) &\geq V_f(\mu'(f)) - \sum_{w \in \mu'(f)} u'(w) \end{aligned}$$

with one inequality strict. Adding these inequalities contradicts (1) for f. Hence (2) $\Rightarrow$ (1).

(1) $\Rightarrow$ (2). Suppose (1) does not hold for firm f. That is,

$$V_f(\mu(f)) - \sum_{w \in \mu(f)} u'(w) > V_f(\mu'(f)) - \sum_{w \in \mu'(f)} u'(w) \tag{9}$$

Let $(\mu, \hat{u}) \in \mathcal{I}_\mu$ (such an outcome exists since μ is IR-feasible). Define the outcome (μ, u) as

[15] Even though $V_{f'}$ may take negative values, in an efficient matching $\mu^\star$, we must have $V_{f'}(\mu^\star(f')) \geq 0$ – else not matching any worker to f' improves aggregate surplus.

follows

$$u(w) = \begin{cases} u'(w) & \text{if } w \in \mu(f) \\ \hat{u}(w) & \text{otherwise} \end{cases}$$

We first show that $(\mu, u) \in \mathcal{I}_\mu$. Since $(\mu', u'), (\mu, \hat{u}) \in \mathcal{I}$, $u(w) \geq 0$ for all $w \in W$ and for all $f' \neq f$,

$$0 \leq V_{f'}(\mu(f')) - \sum_{w \in \mu(f')} \hat{u}(w) = V_{f'}(\mu(f')) - \sum_{w \in \mu(f')} u(w)$$

where the inequality follows from $(\mu, \hat{u}) \in \mathcal{I}$ and the equality follows from definition of u. For firm f,

$$V_f(\mu(f)) - \sum_{w \in \mu(f)} u(w) = V_f(\mu(f)) - \sum_{w \in \mu(f)} u'(w) > V_f(\mu'(f)) - \sum_{w \in \mu'(f)} u'(w) \geq 0$$

where the strict inequality is due to (9) and last inequality is true because $(\mu', u') \in \mathcal{I}$. Hence, $(\mu, u) \in \mathcal{I}_\mu$.

Further, by construction of u, $u(w) = u'(w)$ for all $w \in \mu(f)$ and (9) ensures that

$$V_f(\mu(f)) - \sum_{w \in \mu(f)} u(w) > V_f(\mu'(f)) - \sum_{w \in \mu'(f)} u'(w)$$

Thus, $(\mu, u) \xrightarrow{f} (\mu', u')$. Hence, (1) $\Rightarrow$ (2). ■

Proof of Lemma 4. Suppose there exists a flow ϕ in the directed graph. Fix $w \in W$. Suppose there is an incoming edge from $\hat{f}$ (i.e., $f \equiv \mu(w)$) and an outgoing edge to $f' \equiv \mu'(w)$ at w, then flow balancing will imply $\phi(w, f') = \phi(\hat{f}, w)$, and then set $u'(w) = \phi(w, f') = \phi(\hat{f}, w)$.

We show that this u' satisfies (**ES**) and (**IR**). To see this, pick $f \in F$. The inflow to vertex f is equal to the flow from f to $\hat{f}$ (since $b(f) = 0$ and flow balancing must hold), and inflow (and outflow) is equal to

$$\sum_{(w,f) \in E} \phi(w, f) = \sum_{w \in \mu'(f)} u'(w) \leq c(f, \hat{f}) = V_f(\mu'(f))$$

where the inequality follows because flow satisfies capacity constraint. Hence, (**IR**) holds.

Now, consider vertex $\hat{f}$ and flow balancing at this vertex. Total inflow is equal to sum of inflows from f and from v_0. Flow from f to $\hat{f}$ is equal to the inflow to vertex f, which in turn is equal to $\sum_{w\in\mu'(f)} u'(w)$. Hence, total inflow to $\hat{f}$ is equal to

$$\sum_{w\in\mu'(f)} u'(w) + \phi(v_0, \hat{f})$$

Total outflow from $\hat{f}$ is

$$\sum_{(\hat{f},w)\in E} \phi(\hat{f}, w) = \sum_{w\in\mu(f)} \phi(\hat{f}, w) = \sum_{w\in\mu(f)} u'(w)$$

Using the fact that $b(\hat{f}) = V_f(\mu'(f)) - V_f(\mu(f))$, flow balancing at $\hat{f}$ gives us

$$\sum_{w\in\mu'(f)} u'(w) + \phi(v_0, \hat{f}) - \sum_{w\in\mu(f)} u'(w) = V_f(\mu'(f)) - V_f(\mu(f))$$

Since $\phi(v_0, \hat{f}) \geq 0$, we get

$$\sum_{w\in\mu'(f)} u'(w) - \sum_{w\in\mu(f)} u'(w) \leq V_f(\mu'(f)) - V_f(\mu(f))$$

Hence, (**ES**) holds for firm f.

The converse proof follows similar arguments. Suppose u' is a feasible solution to (**ES**) and (**IR**). Then, we can construct a flow in $\mathcal{G}$ from it. For every $w \in W$, since $\mu'(w) \neq \emptyset$ and $\mu(w) \neq \emptyset$, we can set $\phi(w, \mu'(w)) = \phi(\hat{f}, w) := u'(w)$, where $\mu(w) = f$. This balances flow at w. Now, at any $f \in F$, let

$$\phi(f, \hat{f}) := \sum_{w\in\mu'(f)} u'(w) = \sum_{w\in\mu'(f)} \phi(w, f)$$

This balances flow at f. Further, since u' satisfies (**IR**), this means

$$\phi(f, \hat{f}) = \sum_{w\in\mu'(f)} u'(w) \leq V_f(\mu'(f)) = c(f, \hat{f})$$

Hence, capacity constraint on edge $(f, \hat{f})$ holds.[16]

Now to define the flow $(v_0, \hat{f})$, let

$$\phi(v_0, \hat{f}) := \big[V_f(\mu'(f)) - V_f(\mu(f))\big] - \big[\sum_{w \in \mu'(f)} u'(w) - \sum_{w \in \mu(f)} u'(w)\big]$$

Since u' satisfies (**ES**), this is non-negative. Also,

$$\phi(v_0, \hat{f}) \leq V_f(\mu'(f)) + \sum_{w \in \mu(f) \setminus \mu'(f)} u'(w) \leq V_f(\mu'(f)) + \sum_{f' \neq f} V_{f'}(\mu'(f')) \leq V(\mu') < c(v_0, \hat{f}),$$

where the second inequality uses the fact that (μ', u') is an IR outcome and the last inequality uses the fact that the capacity of edge $(v_0, \hat{f})$ is sufficiently large. Hence, capacity constraint of edge $(v_0, \hat{f})$ holds for each $f \in F$.

Next,

$$\phi(v_0, \hat{f}) = \big[V_f(\mu'(f)) - V_f(\mu(f))\big] - \big[\sum_{w \in \mu'(f)} u'(w) - \sum_{w \in \mu(f)} u'(w)\big] = b(\hat{f}) - [\phi(f, \hat{f}) - \sum_{w \in \mu(f)} \phi(\hat{f}, w)]$$
$$\implies \phi(v_0, \hat{f}) + \phi(f, \hat{f}) - \sum_{w \in \mu(f)} \phi(\hat{f}, w) = b(\hat{f})$$

This means flow balancing holds at $\hat{f}$.

Finally, at vertex v_0,

$$\sum_{f \in F} \phi(v_0, \hat{f}) = \sum_{f \in F} \big[V_f(\mu'(f)) - V_f(\mu(f))\big] - \sum_{f \in F} \big[\sum_{w \in \mu'(f)} u'(w) - \sum_{w \in \mu(f)} u'(w)\big]$$
$$= \sum_{f \in F} \big[V_f(\mu'(f)) - V_f(\mu(f))\big] = -b(v_0)$$

where the second equality follows because all workers are matched in μ' and μ, and hence,

$$\sum_{f \in F} \big[\sum_{w \in \mu'(f)} u'(w) - \sum_{w \in \mu(f)} u'(w)\big] = 0$$

Since there are no incoming edges to v_0, this shows that flow balancing holds at v_0 also. ■

[16] Capacity of other edges is greater than $V(\mu')$ and $u'(w) \leq V_{\mu'(w)}(\mu'(\mu'(w))) \leq V(\mu')$ implies capacity constraints of edges $(w, \mu'(w))$ and $(\hat{f}, w)$, where $\mu(w) = f$, also hold.

C A vNM stable set with only inefficient outcomes

Example 4 (a vNM stable set with only inefficient outcomes)**.** Consider the following economy. There are three firms $F = \{f_1, f_2, f_3\}$ and three workers $W = \{w_1, w_2, w_3\}$. The surplus generated by each firm is summarized in Table 2.

	$\{w_1\}$	$\{w_2\}$	$\{w_3\}$	$\{w_1, w_2\}$	$\{w_2, w_3\}$	$\{w_3, w_1\}$	$\{w_1, w_2, w_3\}$
$V_{f_1}(\cdot)$	-1	-1	-1	4	-1	2	-1
$V_{f_2}(\cdot)$	3	3	1	-1	-1	-1	-1
$V_{f_3}(\cdot)$	-1	-1	2	-1	2	-1	-1

Table 2: Surplus of firms and workers

There is a unique efficient matching $\mu^\star$ given by $\mu^\star(f_1) = \{w_1, w_2\}$, $\mu^\star(f_2) = \emptyset$, and $\mu^\star(f_3) = \{w_3\}$. The core is empty for this economy.

Consider the following three inefficient matchings:

$$\begin{array}{lccc} \mu_1: & \mu_1(f_1) = \{w_1, w_3\} & \mu_1(f_2) = \{w_2\} & \mu_1(f_3) = \emptyset \\ \mu_2: & \mu_2(f_1) = \{w_1, w_2\} & \mu_2(f_2) = \{w_3\} & \mu_2(f_3) = \emptyset \\ \mu_3: & \mu_3(f_1) = \emptyset & \mu_3(f_2) = \{w_1\} & \mu_3(f_3) = \{w_2, w_3\} \end{array}$$

Let $\bar{u}$ be given by $\bar{u}(w_1) = \bar{u}(w_2) = 2$ and $\bar{u}(w_3) = 0$. For every $i \in \{1, 2, 3\}$, define the outcome $a_i = (\mu_i, \bar{u})$. At each of the outcomes a_1, a_2, and a_3, firms f_1 and f_3 receive payoff 0, while firm f_2 receives payoff 1. Moreover, workers w_1 and w_2 receive payoff 2, while worker w_3 receives payoff 0. Now define the set of outcomes

$$\mathcal{U} = \{(\mu_1, u) \in \mathcal{I}_{\mu_1} : u(w_3) > 0,\ u(w_1) + u(w_3) < 2,\ u(w_2) < 2\}.$$

Let $\mathcal{V} = \mathcal{U} \cup \{a_1, a_2, a_3\}$. By construction, $\mathcal{V} \subseteq \mathcal{I}$. We show that $\mathcal{V}$ is a vNM stable set.

Internal stability: By Observation 1, no two outcomes in $\mathcal{U} \cup \{a_1\}$ can dominate one another. Moreover, the payoff of each worker and firm is the same at a_1, a_2 and a_3. Hence, none of these outcomes dominates another. It remains to compare outcomes in $\mathcal{U}$ with a_2 and a_3. Consider any $(\mu_1, u) \in \mathcal{U}$, and let $u(w_1) = x$, $u(w_2) = y$, and $u(w_3) = z$. By the definition of $\mathcal{U}$, $z > 0$, $x + z < 2$, and $y < 2$. Hence, the payoffs of firms f_1, f_2 and f_3 at (μ_1, u) are $2 - x - z > 0$, $3 - y > 1$ and 0, respectively.

Since a_1, a_2, and a_3 give the firms and workers the same payoff, if (μ_1, u) dominates either

a_2 or a_3, it would also dominate a_1. This contradicts Observation 1. Hence, $(\mu_1, u) \nrightarrow a_2$ and $(\mu_1, u) \nrightarrow a_3$.

At a_2 and a_3, firm f_1 receives payoff 0, so neither outcome can dominate (μ_1, u) through f_1. At a_3, firm f_2 receives payoff 1, whereas its payoff at (μ_1, u) is strictly greater than 1. Hence, a_3 cannot dominate (μ_1, u) through f_2. At a_2, firm f_2 is matched to w_3, but w_3 receives payoff 0 at a_2 and $z > 0$ at (μ_1, u). Thus, a_2 cannot dominate (μ_1, u) through f_2. At a_2 and (μ_1, u), f_3 is unmatched and receives payoff 0 at both outcomes, so there is no strict improvement. At a_3, f_3 is matched to $\{w_2, w_3\}$, but w_3 receives 0 at a_3 and $z > 0$ at (μ_1, u). Hence, $a_2 \nrightarrow (\mu_1, u)$ and $a_3 \nrightarrow (\mu_1, u)$. Thus, no two outcomes in $\mathcal{V}$ dominate one another, and $\mathcal{V}$ satisfies internal stability.

External stability: Consider any $(\mu, u) \in \mathcal{I} \setminus \mathcal{V}$. Since (μ, u) is individually rational, f_2 can only be matched to $\emptyset$, $\{w_1\}$, $\{w_2\}$, or $\{w_3\}$. We consider these four cases.

Case 1: $\mu(f_2) = \emptyset$.
In this case the payoff of f_2 is zero. If $u(w_2) \leq 2$, then $a_1 \xrightarrow{f_2} (\mu, u)$. If $u(w_2) > 2$, then w_2 cannot be unmatched. Moreover, since $(\mu, u) \in \mathcal{I}$ and $\mu(f_2) = \emptyset$, we must have $\mu(f_1) = \{w_1, w_2\}$ ($\mu(f_3) = \{w_2, w_3\}$ is not possible since $V_{f_3}(\{w_2, w_3\}) = 2 < u(w_2)$, which will contradict IR for firm f_3). Since $V_{f_1}(\{w_1, w_2\}) = 4$ and $u(w_2) > 2$, it follows that $u(w_1) < 2$. Hence, $a_3 \xrightarrow{f_2} (\mu, u)$.

Case 2: $\mu(f_2) = \{w_1\}$.
Since $(\mu, u) \in \mathcal{I}$, we must have $\mu(f_1) = \emptyset$ and $u(w_2) \leq 2$. If $u(w_1) < 2$, then $a_2 \xrightarrow{f_1} (\mu, u)$. If $u(w_1) > 2$, then the payoff of f_2 is strictly less than 1. Combining this with the fact that $u(w_2) \leq 2$, $a_1 \xrightarrow{f_2} (\mu, u)$. If $u(w_1) = 2$, the payoff of f_2 is 1. Hence, if $u(w_2) < 2$, then $a_1 \xrightarrow{f_2} (\mu, u)$. However, if $u(w_2) = 2$, by IR of (μ, u), $\mu(f_3) = \{w_2, w_3\}$ and $u(w_3) = 0$. But, then $(\mu, u) = a_3$, contradicting $(\mu, u) \notin \mathcal{V}$.

Case 3: $\mu(f_2) = \{w_2\}$.
Since $(\mu, u) \in \mathcal{I}$, in this case either $\mu(f_1) = \emptyset$ or $\mu(f_1) = \{w_1, w_3\}$. Suppose $\mu(f_1) = \emptyset$. Then, the payoff of f_1 is zero. Since w_1 cannot be matched to f_3 at an IR outcome, w_1 is unmatched and hence $u(w_1) = 0$. If $u(w_2) < 2$, then $a_2 \xrightarrow{f_1} (\mu, u)$. If $u(w_2) \geq 2$, then the payoff of f_2 is weakly less than 1. Hence, $a_3 \xrightarrow{f_2} (\mu, u)$. Now suppose $\mu(f_1) = \{w_1, w_3\}$. By IR of (μ, u), $u(w_1) + u(w_3) \leq 2$. If $u(w_2) > 2$, then $a_3 \xrightarrow{f_2} (\mu, u)$. Suppose $u(w_2) = 2$. If $u(w_1) < 2$, then $a_3 \xrightarrow{f_2} (\mu, u)$. If $u(w_1) = 2$, then IR implies $u(w_3) = 0$, so $(\mu, u) = a_1$, contradicting $(\mu, u) \notin \mathcal{V}$. Finally, suppose $u(w_2) < 2$. If $u(w_3) = 0$, then $a_3 \xrightarrow{f_3} (\mu, u)$. If

$u(w_3) > 0$, then, since $(\mu, u) \notin \mathcal{U}$, we must have $u(w_1) + u(w_3) = 2$. Hence, $a_2 \xrightarrow{f_1} (\mu, u)$.

Case 4: $\mu(f_2) = \{w_3\}$.
Since $(\mu, u) \in \mathcal{I}$, in this case $\mu(f_3) = \emptyset$ and either $\mu(f_1) = \emptyset$ or $\mu(f_1) = \{w_1, w_2\}$. Moreover, $u(w_3) \leq 1$ and hence the payoff of f_2 is at most 1. If $\mu(f_1) = \emptyset$, then $a_2 \xrightarrow{f_1} (\mu, u)$. Suppose $\mu(f_1) = \{w_1, w_2\}$. Since $V_{f_1}(\{w_1, w_2\}) = 4$, IR implies $u(w_1) + u(w_2) \leq 4$. If $u(w_1) < 2$, then $a_3 \xrightarrow{f_2} (\mu, u)$ and if $u(w_2) < 2$, then $a_1 \xrightarrow{f_2} (\mu, u)$. The only remaining case is $u(w_1) = u(w_2) = 2$. If $u(w_3) > 0$, then $a_1 \xrightarrow{f_2} (\mu, u)$ and $a_3 \xrightarrow{f_2} (\mu, u)$. If $u(w_3) = 0$, then $(\mu, u) = a_2$, contradicting $(\mu, u) \notin \mathcal{V}$.

Thus, every outcome in $\mathcal{I} \setminus \mathcal{V}$ is dominated by an outcome in $\mathcal{V}$. Hence, $\mathcal{V}$ satisfies external stability.