

Gersbach, Hans; Haller, Hans

Working Paper

Matching on Bipartite Graphs

CESifo Working Paper, No. 5575

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Gersbach, Hans; Haller, Hans (2015) : Matching on Bipartite Graphs, CESifo Working Paper, No. 5575, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/123215>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Working Papers

www.cesifo.org/wp

Matching on Bipartite Graphs

Hans Gersbach
Hans Haller

CESIFO WORKING PAPER NO. 5575

CATEGORY 4: LABOUR MARKETS

OCTOBER 2015

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp

ISSN 2364-1428

CESifo

Center for Economic Studies & Ifo Institute

Matching on Bipartite Graphs

Abstract

We study stable matchings on exogenously given or endogenously formed bipartite graphs that reflect constraints on matching. The effects of a change of the exogenously given graph (constraints) are investigated. Specifically, we examine to what extent individuals gain or lose from relaxing restrictions on matching. Sabotage of particular matches or location choice can render certain matches infeasible and, therefore, lead to endogenous constraints.

JEL-Codes: C780, D020, D850.

Keywords: matching theory, networks, graphs.

Hans Gersbach
CER-ETH – Center of Economic
Research at ETH Zurich
Zürichbergstrasse 18
Switzerland – 8092 Zurich
hgersbach@ethz.ch

Hans Haller
Department of Economics
Virginia Polytechnic Institute and
State University
USA – Blacksburg, VA 24061-0316
haller@vt.edu

This Version: September 2015

We would like to thank Mario Gersbach, Jonathan Lorand and Simon Skok for helpful comments.

1 Introduction

Motivation

Often pairwise matching is not feasible although it might be desirable. Typical examples can be found in the market for top managers and firms and in the marriage market. In the former case, particular matches may not be feasible as the manager has worked recently for a competitor or may live geographically too far away.¹ In the latter case, segregation of various sorts, be it racial, religious, political or geographical can make matches infeasible.

Constraints on feasible matches can be given or may be altered. Three processes that affect the feasibility of matches are particularly important. First, segregation may simply disappear or become less binding. For instance, globalization does not only entail integration of product markets, but also triggers integration of specific subsegments of labor markets such as the market for top managers² or for scientists. The integration of such markets means that more matches become feasible. Second, people may take decisions to alter the set of possible matches. This is most obvious in the marriage market when individuals change their location.

Third, individuals may engage in activities to render a particular match of other individuals infeasible – possibly to their own advantage. Such sabotage activities may take several forms. Recommendations or contract clauses developed by firms whose CEO has left could be used to deter the hiring of the CEO by a particular competitor. A company may actively try to convince a customer to cancel an order for an investment good at a rival company.³ In marketing campaigns firms may stress that a rival is unattractive for particular customers in order to deter these matches (see, e.g., Lubin (2011) and Chakrabarti and Haller (2011)). In the marriage market, avoiding meetings or locations which could otherwise bring people together, or even spreading particular rumors are part of a wide range of activities aimed at rendering certain matches infeasible. Social media apparently have enhanced those possibilities.

¹In some of our analysis, we consider instances where geographical boundaries cease to constrain such matches.

²See Gersbach and Schmutzler (2014) and the references therein.

³See Friedman (1998), p. 577, for a specific example.

Approach

In this paper, we take first steps towards addressing the consequences of feasibility constraints on two-sided matching, when these constraints are given or can be moved – weakened in the case of desegregation or tightened in the case of sabotage.

In the first part of the paper, we are particularly interested in how stable matchings are affected by feasibility constraints, and whether and which individuals gain when the set of feasible matches is expanded. In the second part of the paper, we study matching when feasibility constraints are endogenous, result from individual decisions. We explore two main processes outlined above. First, we investigate whether and how individuals gain from sabotaging feasible matches and how restrictive feasibility constraints become when individuals can sabotage other matches. Second, we study matching cum location choices where the location choices determine the feasible matches. Here, we are particularly interested in the existence and Pareto optimality of stable matchings of persons and locations.

The focus of our paper lies on two-sided pairwise matching as in the seminal contribution of Gale and Shapley (1962). In the marriage market interpretation, the population consists of men and women. A matching selects heterosexual couples such that each individual is matched with exactly one partner of the other sex or remains unmatched. Stability requires that no matched person prefers to be single and no pair consisting of a man and a woman prefers being a couple to the status quo. This presupposes that individuals have preferences over partners, including having no partner. In the standard setting of two-sided matching, in principle, a man can be matched with any woman and vice versa – if one disregards preferences.

To study the consequences of feasibility constraints on matchings, we employ a different notion of two-sided pairwise matching that is based on bipartite graphs. Again the population (which also constitutes the set of nodes or vertices) is partitioned into two sets, men and women for our specific application. An edge connects a man and a woman. Formally, it is the unordered pair consisting of the particular man and woman. Taking the set of men and the set of women as given, a bipartite graph is identified by its set of edges. A matching on the graph is a subset of its edges, no two of which share an endpoint.⁴ We are considering two-sided

⁴For an introduction to matching on (not necessarily bipartite) graphs, see Lovász and Plummer (1986).

matching à la Gale and Shapley when not all matches are feasible even if they would be desirable. Feasibility will be described by a **bipartite graph** (**bigraph** for short) whose edges represent the feasible matches. Like in Gale and Shapley (1962), individuals have preferences over potential partners, including having no partner. In other words, we combine matching on a bigraph with the stability requirements of Gale and Shapley.

Results

For the first part of the paper, we do not have to be concerned about existence of stable matchings, given some constraints on the feasible matches. Indeed we shall show below, as an immediate consequence of the classical result of Gale and Shapley (1962), that a stable matching exists for any combination of bigraph and complete and transitive preferences. The main welfare implications in Sections 3 and 4 of the paper are as follows:

Obviously, if the feasibility constraints for matches are relaxed, that is, if there are more matching opportunities, then a stable status quo matching need no longer be stable. We find that if, indeed, the status quo is destabilized, then in every ensuing stable matching under the relaxed feasibility constraints, at least one individual is better off than at the status quo. More specifically, if a couple blocks the status quo, then at least one of the partners will fare better; moreover, in some stable matching under the relaxed feasibility constraints, none of the partners will fare worse.

We establish a general result that the number of losers cannot exceed the number of winners when feasibility constraints for matches are relaxed. Relaxation of feasibility constraints can affect all individuals, but the number of losers can at most be as large as the number of winners.

In Section 5, we leave the world of minor changes with potentially drastic consequences and turn to the effects of a large expansion of matching opportunities. Suppose that two identical hitherto segregated marriage markets become desegregated. Will the stable matchings for the segregated markets give rise to a stable matching in the desegregated market? The answer is clearly no if individuals have a strict preference for partners from the “other” subpopulation, partners who were unavailable under segregation. The answer is clearly yes if individuals have a strict

preference for partners from the “own” subpopulation, partners who were already available under segregation. In addition to these polar cases, many intermediate scenarios exist, in particular the following: Let us label the segregated markets A and B . Let us assume interchangeability: A person is indifferent between being matched with an A-partner and being matched with the corresponding B-partner. Suppose the status quo consists of the men-optimal stable matching in market A and the women-optimal stable matching in market B . Then in the desegregated market a man m and a woman w can block the status quo only if m belongs to B and w belongs to A , that is if each has been in the worst match possible among all stable matchings of the respective segregated market.

Drawing from our general result regarding winners and losers we observe a stark contrast between economic integration and integration of marriage markets. When originally separated pure exchange economies are integrated, almost all consumers may lose in some instances. The first welfare theorem only guarantees that at least one consumer gains. While they can also be gainers and losers from the integration of marriage markets, there cannot be more losers than gainers.

Part 2 deals with endogenous constraints on matching and begins with Section 6 where sabotage by women is investigated. In general, refraining from sabotage is not in the interest of all women even if we consider the women-optimal stable matching. Women may lose through sabotaging even if we start with the men-optimal stable matching. However, it is possible that a stable matching that would have prevailed without sabotage, reemerges as a Nash sabotage outcome. It is also worth noting that a woman may gain from sabotaging (preventing) her own match in the women maximal matching.

In Section 7, location choice means endogenous segregation. In case individuals have preferences for partners only, mobile individuals choose the location of the partner with whom they prefer to be matched, provided this is possible. In case individuals have preferences for partners and locations, we show existence of a stable matching and weak core equivalence.

Outline

After a short elaboration on the marriage theorem, the next section reviews the two basic concepts of matching, matching on bipartite graphs and pairwise match-

ing in the tradition of Gale and Shapley (1962). In Section 3, we introduce a model that incorporates both concepts, where the bipartite graph (bigraph) reflects constraints on matching. We show existence of stable matchings and perform some initial welfare analysis. Section 4 explores the effects of small changes in matching opportunities. Section 5 deals with some large changes in matching opportunities. Section 6 is devoted to sabotage. In Sections 7 and 8, we consider the choice of locations and partners. Section 9 concludes.

2 Preliminaries

2.1 Bipartite Graphs and Matching

We consider a finite population I which constitutes the set of nodes or vertices of the graphs to be investigated. The population I is partitioned into two non-empty subsets M and W , with $|M| = k$ and $|W| = \ell$. The subpopulation M stands for men (males). The subpopulation W stands for women (females). A bipartite graph or bigraph based on the given partition is an undirected graph whose edges are of the form $\{m, w\}$ for some $m \in M$ and some $w \in W$. For given sets I, M and W , a graph $G = (I, \mathcal{E})$ is identified by the set of its edges, $\mathcal{E}$. We simply refer to the graph $\mathcal{E}$ if this causes no confusion. Depending on the application, $\{m, w\} \in \mathcal{E}$ can have different meanings. For instance, it could mean that m and w are mutually acceptable as partners. Or it could mean that they are acquaintances, etc. The complete bigraph based on M and W is defined as $\mathcal{C} = \{\{m, w\} \mid (m, w) \in M \times W\}$. Then a bigraph given M and W is a subset $\mathcal{E}$ of $\mathcal{C}$.

A matching on a graph $\mathcal{E}$ is a subset $\mathcal{M}$ of its edges, no two of which share an element. That is, $\mathcal{M} \subseteq \mathcal{E}$ and $\{m, w\} \cap \{m', w'\} = \emptyset$ for $\{m, w\}, \{m', w'\} \in \mathcal{M}$, $\{m, w\} \neq \{m', w'\}$. Perhaps the first formal treatment of marriage and matching is found in a combinatorial lemma known as the “marriage theorem”, also known as Hall’s theorem. It was first stated and shown in its definitive form in Hall (1935) and Maak (1936). Weyl (1949) introduced the term marriage theorem. Jacobs (1969), pp. 105-106 provides a proof and statement of the theorem. We state it in slightly different form:

Theorem 1 (Marriage Theorem). *Let M be a finite non-empty set, W be a non-empty (possibly infinite) set, and $D : M \rightarrow W$ be a correspondence. Define $D(N) = \bigcup_{m \in N} D(m)$ for $N \subseteq M$. Then there exists an injective selection d of D if and only if $|D(N)| \geq |N|$ for all $N \subseteq M$.*

The standard interpretation takes M as set of men, W as set of women, and $D(m)$ as the set of m 's acquaintances. In an injective selection d of D , each man is matched with exactly one acquaintance and each woman is matched with at most one man. There are conceivable alternative interpretations of a set $D(m)$ as the set of women whom m considers acceptable partners or as the set of women who find m acceptable. Also, $D(m)$ could stand for the set of women whom m considers acceptable and who consider m acceptable. The correspondence D can be identified with a bipartite graph $\mathcal{E}$ as follows: $\{m, w\} \in \mathcal{E} \Leftrightarrow w \in D(m)$. The injective selection d defines a matching $\mathcal{M}$ in $\mathcal{E}$: $\{m, w\} \in \mathcal{M} \Leftrightarrow w = d(m)$. In case $|M| = |W|$, the marriage theorem provides a necessary and sufficient condition for a perfect matching, that is a matching where every member of $I = M \cup W$ is matched with an acceptable partner of the opposite sex.

2.2 Stable Matchings

The fact that $\{m, w\} \in \mathcal{E}$ and $\{m, w'\} \in \mathcal{E}$ tells us that both $\{m, w\}$ and $\{m, w'\}$ are acceptable (or feasible) matches, but does not convey any information which one m prefers. In contrast, the two-sided matching model of Gale and Shapley assumes complete and transitive preferences for partners. Some of the subsequent notation and terminology is adopted from Roth and Sotomayor (1990). There is again a finite population I that is partitioned into a non-empty male subpopulation $M = \{m_1, \dots, m_k\}$ and a non-empty female subpopulation $W = \{w_1, \dots, w_\ell\}$. A **matching** is a bigraph $\mathcal{M}$ based on the given partition such that no two of its edges share an element: $\{m, w\} \cap \{m', w'\} = \emptyset$ for $\{m, w\}, \{m', w'\} \in \mathcal{M}$, $\{m, w\} \neq \{m', w'\}$. No further a priori restrictions are imposed.

Moreover, individuals have preferences for partners, including the possibility of not having a partner. Each $m \in M$ has complete and transitive preferences on $W \cup \{m\}$, represented by an ordered list $P(m)$. For example

$$P(m) = w_1, w_2, m, w_3, \dots, w_\ell$$

means that m 's first choice is to be married to woman w_1 , his second choice is to be married to woman w_2 , and his third choice is to remain single. Women w_3 to w_ℓ are not acceptable to m . He would rather be single than be married to one of them. An abbreviated list contains only the acceptable women:

$$P(m) = w_1, w_2.$$

Similarly, each $w \in W$ has complete and transitive preferences on $M \cup \{w\}$, represented by an ordered list $P(w)$. If a person is indifferent between several possible mates, then those are put in square brackets in the preference list:

$$P(w) = m_2, [m_3, m_4, w], m_1$$

means that m_2 is w 's first choice; to be married to m_3 , to be married to m_4 and being single tie for second choice while m_1 is her third choice. The abbreviated list is

$$P(w) = m_2, [m_3, m_4, w].$$

$\mathbf{P} = (P(m_1), \dots, P(m_k), P(w_1), \dots, P(w_\ell))$ denotes a preference profile, that is a collection of lists, one for each individual. Then a particular marriage market is specified by a triple $(M, W; \mathbf{P})$. The notation $w >_m w'$ means that m prefers woman w to woman w' and $w \geq_m w'$ means that m likes w at least as much as w' . $m >_w m'$ and $m \geq_w m'$ are defined in an analogous way. Woman w is acceptable to man m if he likes her at least as much as remaining single, i.e., $w \geq_m m$. Analogously, m is acceptable to w if $m \geq_w w$. An individual is said to have strict preferences if he or she is not indifferent between any two acceptable alternatives.

Following Holzman and Samet (2014), we say that **men are universally ranked** if they are ranked the same way by all women, that is $P(w_1) = P(w_2) = \dots = P(w_\ell)$ and similarly, **women are universally ranked** if they are ranked the same way by all men, that is $P(m_1) = P(m_2) = \dots = P(m_k)$.

A matching $\mathcal{M}$ can be identified with a bijection $\mu : I \rightarrow I$ with the properties that (a) $\mu^{-1} = \mu$; (b) $\mu(m) \notin W$ implies $\mu(m) = m$ for $m \in M$; (c) $\mu(w) \notin M$ implies $\mu(w) = w$ for $w \in W$. Namely, set $\mu(m) = w$ and $\mu(w) = m$ if $\{m, w\} \in \mathcal{M}$. Set $\mu(i) = i$ if there is no j such that $\{i, j\} \in \mathcal{M}$. Conversely, for $m \in M, w \in W$, $\{m, w\}$ belongs to $\mathcal{M}$ if and only if $m = \mu(w)$.

Definition 1. A matching $\mathcal{M}$ (or, equivalently, μ) in the marriage market $(M, W; \mathbf{P})$

is stable if:

- (i) For any $m \in M$, $w \in W$ who are matched in $\mathcal{M}$, neither partner wants to go single. That is, $\{m, w\} \in \mathcal{M}$ (or, equivalently $m = \mu(w)$) implies $w \geq_m m$ and $m \geq_w w$.
- (ii) There is no pair $(m, w) \in M \times W$ who both prefer getting married to each other to the status quo. That is, $\mu(m) \geq_m w$ or $\mu(w) \geq_w m$ for all $m \in M$, $w \in W$.

Regardless of the numbers of men and women and their preferences, a stable matching always exists.

Proposition 1 (Gale and Shapley (1962)).

Every marriage market $(M, W; \mathbf{P})$ has a stable matching.

There may be several stable matchings. But sometimes, a unique stable matching exists. For later reference, let us consider such a special case.

Example 1. Suppose that $k = \ell$, all women are acceptable to all men, all men are acceptable to all women, all men have identical strict preferences, and all women have identical strict preferences. Without restriction, we assume that

$$P(m) = w_1, \dots, w_k \text{ for all } m \in M,$$

$$P(w) = m_1, \dots, m_k \text{ for all } w \in W.$$

□ □

For this example, we obtain:

Fact 1. *The unique stable matching is given by $\mu(m_i) = w_i$ for $i = 1, \dots, k$.*

Proof. First of all, a stable matching exists, by Proposition 1. Second, since any man is acceptable to any woman and vice versa, in a stable matching every person is married to an individual of the other sex.

Now let μ be a stable matching. If $\mu(m_1) \neq w_1$, then $\mu(w_1) \neq m_1$ and $w_1 >_{m_1} \mu(m_1)$, $m_1 >_{w_1} \mu(w_1)$; consequently, m_1 and w_1 both prefer to be married to each other to the status quo, contradicting the stability of μ . Hence $\mu(m_1) = w_1$ (and $\mu(w_1) = m_1$). But then, by the same argument, $\mu(m_2) = w_2$ and $\mu(w_2) = m_2$. Iteration of the argument yields $\mu(m_i) = w_i$ for $i = 1, \dots, k$. ■

Several remarks are in order: The unique stable matching in this example exhibits positive assortative matching in the sense that more attractive males marry more attractive women and vice versa. Holzman and Samet (2014) observe that there is still a unique stable matching — which may not be assortative — if $k = \ell$, all preferences are strict and one side of the marriage market is universally ranked, say $P(w) = m_1, m_2, \dots, m_k$ for all $w \in W$. In that case, m_1 will be matched with his top choice, m_2 will be matched with his best choice among the remaining $k - 1$ women, etc.⁵

On the other hand, positive assortative matching is preserved while uniqueness may be lost when each indifference class for men corresponds to an equally sized indifference class for women, that is

$$\begin{aligned} M &= \{m_{11}, m_{12}, \dots, m_{1n_1}, m_{21}, \dots, m_{2n_2}, \dots, m_{k1}, \dots, m_{kn_k}\}, \\ W &= \{w_{11}, w_{12}, \dots, w_{1n_1}, w_{21}, \dots, w_{2n_2}, \dots, w_{k1}, \dots, w_{kn_k}\}, \\ P(m) &= [w_{11}, w_{12}, \dots, w_{1n_1}], [w_{21}, \dots, w_{2n_2}], \dots, [w_{k1}, \dots, w_{kn_k}] \forall m \in M, \\ P(w) &= [m_{11}, m_{12}, \dots, m_{1n_1}], [m_{21}, \dots, m_{2n_2}], \dots, [m_{k1}, \dots, m_{kn_k}] \forall w \in W. \end{aligned}$$

In case $k \neq \ell$, $P(m) = w_1, \dots, w_\ell$ for all $m \in M$ and $P(w) = m_1, \dots, m_k$ for all $w \in W$, there exists a unique stable matching given by $\mu(m_i) = w_i$ for $i = 1, \dots, \min\{k, \ell\}$.

In general, stable matchings are neither unique nor assortative even if $k = \ell$, all women are acceptable to all men, all men are acceptable to all women, all men have identical preferences, and all women have identical preferences.

The common usage of the term assortative matching refers to endowments broadly defined rather than preferences. Becker (1973) pioneered its formal treatment in an economic context whereas earlier informal discussions appeared in the sociological and economic literature. Becker considers an assignment game where the surplus generated by a couple $\{m, w\}$ is of the form $Z_{mw} = f(x_m, y_w)$ when m is endowed with the amount x_m of a quantitative trait like intelligence, education or height and w is endowed with the amount y_w of another or the same quantitative trait. Positive assortative matching here means that if $\{m, w\}$ and $\{m', w'\}$ belong to a stable matching, then $x_m > x_{m'}$ if and only if $y_w > y_{w'}$. This occurs

⁵These two findings are also stated and shown as Lemmata 1.5.1 and 1.5.2 in Gusfield and Irving (1989).

when there is complementarity between the male and the female trait of the form

$$\frac{\partial^2 f}{\partial x \partial y} > 0,$$

that is f exhibits increasing differences or supermodularity (or strategic complements in game-theoretic terminology). Becker's analysis relies in part on the earlier work of Koopmans and Beckmann (1957). The model of an assignment game developed and studied by Shapley and Shubik (1972) encompasses Becker's as a special case. Becker's additional structure allows to address specific questions such as assortative matching, propensity to marry, propensity to work.

When preferences are strict, then the stable matchings form a lattice with respect to the partial order induced by men's preferences. See Theorem 2.16 in Roth and Sotomayor (1990). In particular, there always exists a unique stable matching μ_M that is best for men, that is $\mu_M(m) \geq_m \mu(m)$ for all stable matchings μ and all men $m \in M$. The men-optimal stable matching μ_M is the worst stable matching for women. There is also a unique best stable matching μ_W for women which is the worst stable matching for men. If the model has at least two stable matchings, then μ_M and μ_W necessarily differ. In the following example, there exist exactly two stable matchings.

Example 2. Let $k = \ell \geq 3$. Consider the following cyclical preferences – with addition and subtraction modulo k .

$$P(m_i) = w_{i+1}, w_{i+2}, \dots, w_{i+k} \text{ for } i = 1, \dots, k;$$

$$P(w_i) = m_i, m_{i-1}, m_{i-2}, \dots, m_{i-(k-1)} = m_{i+1} \text{ for } i = 1, \dots, k.$$

In this example, there is one stable matching μ_M given by $\mu_M(m_i) = w_{i+1}$ for $i = 1, \dots, k$ where each man is matched with his most preferred woman (and each woman is matched with her second choice). There is another stable matching μ_W given by $\mu_W(w_i) = m_i$ for $i = 1, \dots, k$ where each woman is matching with her most preferred man (and each man is matched with his second choice). These are the only stable matchings. To show this, let μ be a matching different from μ_M and μ_W . First, consider the case that in μ , in each matched couple, either the man is matched with his favorite woman or the woman is matched with her favorite man. Without restriction, we may assume $\mu(m_1) = w_2, \dots, \mu(m_i) = w_{i+1}, \mu(m_{i+1}) \neq w_{i+2}$ with $1 \leq i < k$. But if m_{i+1} is not matched with his favorite woman, he must

be the favorite man of his partner $\mu(m_{i+1}) \in \{w_{i+3}, \dots, w_k, w_1\}$. Consequently, $m_{i+1} \in \{m_{i+3}, \dots, m_k, m_1\}$, which is impossible. Hence we can rule out this case. Second, consider the case that in μ , there is a matched couple where neither partner is matched with the favorite partner of the opposite sex. Without restriction, we may assume $\mu(m_1) = w_j$ with $j \in \{3, \dots, k\}$. But then $w_j >_{m_{j-1}} \mu(m_{j-1})$ and $m_{j-1} >_{w_j} m_1 = \mu(w_j)$. This shows that m_{j-1} and w_j prefer each other to their partners in μ . Therefore, μ is unstable. $\square \square$

In Example 2.10 of Roth and Sotomayor (1990), there are two stable matchings, μ_M and μ_W , with $\mu_M(m) >_m \mu_W(m)$ for some but not all $m \in M$.

Part I

Exogenous Feasibility Constraints

3 Stable Matchings on Bipartite Graphs

From now on, we consider marriage markets where individuals have preferences for mates like in the model of Gale and Shapley (1962), but the set of possible marriages is restricted. The primitives of the model are given by a quadruple $(M, W; \mathbf{P}; \mathcal{E})$ where $\{M, W\}$ is a partition of the population I into two non-empty subpopulations. The triple $(M, W; \mathbf{P})$ represents an unrestricted marriage market à la Gale and Shapley as described in 2.2. $\mathcal{E}$ is a bigraph based on M and W as introduced in 2.1. It imposes a restriction on marriages: A marriage or couple is only feasible, can only be formed if the corresponding edge belongs to $\mathcal{E}$. $\mathcal{E} = \mathcal{C}$ yields an unrestricted marriage market as a special case. $\mathcal{E}_\emptyset$ denotes the empty graph, which has zero edges. We define a matching in an analogous way as in 2.1:

Definition 2. *A matching in the marriage market $(M, W; \mathbf{P}; \mathcal{E})$ is a subset $\mathcal{M}$ of $\mathcal{E}$'s edges, no two of which share an element.*

That is, $\mathcal{M} \subseteq \mathcal{E}$ and $\{m, w\} \cap \{m', w'\} = \emptyset$

for $\{m, w\}, \{m', w'\} \in \mathcal{M}$, $\{m, w\} \neq \{m', w'\}$.

Like before, a matching $\mathcal{M}$ can be identified with a bijection $\mu : I \rightarrow I$ with the properties that (a) $\mu^{-1} = \mu$; (b) $\mu(m) \notin W$ implies $\mu(m) = m$ for $m \in M$; (c) $\mu(w) \notin M$ implies $\mu(w) = w$ for $w \in W$. The previous definition of a stable matching has to be slightly modified:

Definition 3. A matching $\mathcal{M}$ (or, equivalently, μ) in the marriage market $(M, W; \mathbf{P}; \mathcal{E})$ is stable if:⁶

- (i) For any $m \in M$, $w \in W$ who are matched in $\mathcal{M}$, neither partner wants to go single. That is, $\{m, w\} \in \mathcal{M}$ (or, equivalently $m = \mu(w)$) implies $w \geq_m m$ and $m \geq_w w$.
- (ii) There is no pair $(m, w) \in M \times W$ who can get married to each other and who prefer the marriage to the status quo. That is, $\mu(m) \geq_m w$ or $\mu(w) \geq_w m$ or $\{m, w\} \notin \mathcal{E}$ for all $m \in M$, $w \in W$.

In particular, a stable matching μ is **individually rational**, that is $\mu(i) \geq_i i$ for all $i \in I$. If a pair (m, w) renders a potential matching unstable, we say that (m, w) blocks or destabilizes the particular matching.

When analyzing the marriage market $(M, W; \mathbf{P}; \mathcal{E})$, one can track the restriction $\mathcal{M} \subseteq \mathcal{E}$ explicitly. Yet instead of treating a couple $\{m, w\} \notin \mathcal{E}$ as infeasible or impossible, one could rather treat it as unacceptable, removing w from the list $P(m)$ and m from the list $P(w)$. After proceeding this way for all $\{m, w\} \in \mathcal{C} \setminus \mathcal{E}$, one obtains a reduced profile $\mathbf{P}^\mathcal{E} = (P^\mathcal{E}(m_1), \dots, P^\mathcal{E}(w_\ell))$. Most often, it proves more convenient to analyze the formally unrestricted model $(M, W; \mathbf{P}^\mathcal{E})$ than the restricted marriage market model $(M, W; \mathbf{P}; \mathcal{E})$.

3.1 Existence

Proposition 1 immediately yields existence of stable matchings for any restricted marriage market $(M, W; \mathbf{P}; \mathcal{E})$.

Proposition 2. Every restricted marriage market $(M, W; \mathbf{P}; \mathcal{E})$ has a stable matching.

Proof. The stable matchings of the restricted marriage market $(M, W; \mathbf{P}; \mathcal{E})$ and of the unrestricted marriage market $(M, W; \mathbf{P}^\mathcal{E})$ coincide. By Proposition 1, $(M, W; \mathbf{P}^\mathcal{E})$ has a stable matching. ■

⁶Diestel (2010), p. 40 gives a definition assuming strict preferences and treating isolated nodes as edges.

Notice that in general, the set of matchings of $(M, W; \mathbf{P}; \mathcal{E})$ and the set of matchings of $(M, W; \mathbf{P}^\mathcal{E})$ do not coincide.

3.2 Welfare Analysis

Here we begin to address how the restriction imposed by a bigraph $\mathcal{E}$ affects the stable matching(s) and, consequently, the welfare of individuals in population $I = M \cup W$. In some instances, such a restriction does not have any impact at all: Let $\mathcal{E}(\mathbf{P}) = \{\{m, w\} \in \mathcal{C} : w \geq_m m, m \geq_w w\}$ denote the set of mutually acceptable matches under the preference profile $\mathbf{P}$. If $\mathcal{E}(\mathbf{P}) \subseteq \mathcal{E}$, then the stable matchings in $(M, W; \mathbf{P})$ and the stable matchings in $(M, W; \mathbf{P}; \mathcal{E})$ coincide. However, as a rule, at least one individual gains from a less restrictive marriage market:

Proposition 3. *Let $(M, W; \mathbf{P}; \mathcal{E})$ and $(M, W; \mathbf{P}; \mathcal{E}')$ be two marriage markets with $\mathcal{E} \subset \mathcal{E}'$. Let μ be a stable matching in $(M, W; \mathbf{P}; \mathcal{E})$. Then either (i) or (ii) holds:*

- (i) μ is a stable matching in $(M, W; \mathbf{P}; \mathcal{E}')$.
- (ii) For every stable matching μ' in $(M, W; \mathbf{P}; \mathcal{E}')$, there exists $i \in I$ with $\mu'(i) >_i \mu(i)$.

Proof. Clearly, if (i) holds, (ii) does not hold: Choose $\mu' = \mu$. Now assume that (ii) does not hold, that is, we can choose a stable matching μ' in $(M, W; \mathbf{P}; \mathcal{E}')$ such that $\mu(i) \geq_i \mu'(i)$ for all $i \in I$. We want to show (i). Since μ is a stable matching in $(M, W; \mathbf{P}; \mathcal{E})$, it is individually rational. Suppose (i) does not hold. Then μ is stable in $(M, W; \mathbf{P}; \mathcal{E})$ and unstable in $(M, W; \mathbf{P}; \mathcal{E}')$. Hence there exist $m \in M$ and $w \in W$ with $w >_m \mu(m), m >_w \mu(w), \{m, w\} \notin \mathcal{E}, \{m, w\} \in \mathcal{E}'$. It follows $w >_m \mu'(m), m >_w \mu'(w)$ and $\{m, w\} \in \mathcal{E}'$. That means that μ' is not stable in $\{m, w\} \in \mathcal{E}'$, a contradiction. Therefore, if (ii) does not hold, (i) has to hold. ■

For instance, let $M = \{m_1, m_2\}$, $W = \{w_1, w_2\}$, $\mathcal{E} = \mathcal{E}_\emptyset$, and $\mathcal{E}' = \mathcal{C}$. Preferences are given by $P(m_1) = w_2, w_1$; $P(w_1) = m_2, m_1$; $P(m_2) = [w_1, m_2]$; $P(w_2) = [m_1, w_2]$. Then $\mu(i) = i$ for all $i \in I$ constitutes the unique stable matching in $(M, W; \mathbf{P}; \mathcal{E})$. μ is not stable in $(M, W; \mathbf{P}; \mathcal{E}')$, since $w_1 >_{m_1} m_1$ and $m_1 >_{w_1} w_1$. There exists a stable matching μ' in $(M, W; \mathbf{P}; \mathcal{E}')$ where m_1 and w_1 are matched

and both m_2 and w_2 remain single. There exists a second stable matching μ'' in $(M, W; \mathbf{P}; \mathcal{E}')$ where the couple that destabilizes μ is not matched; rather m_1 is matched with w_2 and w_1 is matched with m_2 . Both m_1 and w_1 rank μ'' over μ' and μ' over μ .

However, when a couple $\{m, w\} \in \mathcal{E}'$ destabilizes a stable matching μ in $(M, W; \mathbf{P}; \mathcal{E})$, it need not be the case that both m and w are better off at a stable matching μ' in $(M, W; \mathbf{P}; \mathcal{E}')$ than at μ . Consider, for example, the case $M = \{m_1, m_2, m_3\}$, $W = \{w_1, w_2, w_3\}$, $\mathcal{E} = \{\{m_2, w_3\}\}$, $\mathcal{E}' = \mathcal{C}$, and preferences given by $P(m) = w_1, w_2, w_3$ for $m \in M$, $P(w) = m_1, m_2, m_3$ for $w \in W$. In the unique stable matching μ in $(M, W; \mathbf{P}; \mathcal{E})$, m_2 and w_3 are matched with each other whereas all others remain single. Now $w_3 >_{m_1} m_1$ and $m_1 >_{w_3} w_3$ so that μ is not stable in $(M, W; \mathbf{P}; \mathcal{E}')$. Yet in the unique stable matching μ' in $(M, W; \mathbf{P}; \mathcal{E}')$, $\mu'(m_i) = w_i$, hence $\mu'(w_3) <_{w_3} \mu(w_3)$.

The question remains whether both members of a couple that can block a matching μ can do worse in a stable matching μ' than at μ . The answer is provided in the following proposition which states that at least one partner will fare better at μ' than at μ .

Proposition 4. *Let μ be a matching in the marriage market $(M, W; \mathbf{P}; \mathcal{E}')$. If $\{m, w\} \in \mathcal{E}'$ is a possible match that blocks μ , that is*

$$w >_m \mu(m) \text{ and } m >_w \mu(w), \quad (1)$$

then at any stable matching μ' in $(M, W; \mathbf{P}; \mathcal{E}')$,

$$\mu'(m) >_m \mu(m) \text{ or } \mu'(w) >_w \mu(w). \quad (2)$$

Proof. Let $(M, W; \mathbf{P}; \mathcal{E}')$, μ and $\{m, w\}$ be as hypothesized and μ' be a stable matching in $(M, W; \mathbf{P}; \mathcal{E}')$. Suppose $\mu'(m) <_m w$ and $\mu'(w) <_w m$. Then $\{m, w\}$ blocks μ' , in contradiction to μ' being a stable matching in $(M, W; \mathbf{P}; \mathcal{E}')$. Hence to the contrary,

$$\mu'(m) \geq_m w \text{ or } \mu'(w) \geq_w m. \quad (3)$$

(3) and (1) imply (2). ■

Proposition 4 has further implications. First, it yields an alternative proof of Proposition 3. Second, Demange et al. (1987) have shown

Proposition 5. *Let μ be an unstable matching in the marriage market $(M, W; \mathbf{P}; \mathcal{E}')$. Then (a) or (b) holds:*

(a) μ is not individually rational.

(b) There exist a pair $\{m, w\} \in \mathcal{E}'$ blocking μ and a stable matching $\bar{\mu}$ in the marriage market $(M, W; \mathbf{P}; \mathcal{E}')$ such that $\bar{\mu}(m) \geq_m \mu(m)$ and $\bar{\mu}(w) \geq_w \mu(w)$.

Notice that the original result was formulated for an unrestricted marriage market. However, the proposition follows from the fact that the formally unrestricted model $(M, W; \mathbf{P}^{\mathcal{E}'})$ and the restricted marriage market model $(M, W; \mathbf{P}; \mathcal{E}')$ have the same set of stable matchings. Actually, Demange et al. (1987) (and Roth and Sotomayor (1990)) assert that either (a) or (b) holds. Consider, however, the following

Example 3. Let $k = \ell = 3$, $\mathcal{E}' = \mathcal{C}$, $P(m_1) = P(m_2) = w_1, w_2$; $P(w_1) = P(w_2) = m_1, m_2$ whereas m_3 and w_3 prefer to remain single.

Let μ be given by $\mu(m_1) = w_2, \mu(m_2) = w_1, \mu(m_3) = w_3$. Then

(a) μ is not individually rational, since $m_3 >_{m_3} \mu(m_3)$ and $w_3 >_{w_3} \mu(w_3)$.

(b) $\{m_1, w_1\}$ is blocking μ . The unique stable matching in the market $(M, W; \mathbf{P}; \mathcal{E}')$ is $\bar{\mu}$ where $\bar{\mu}(m_1) = w_1, \bar{\mu}(m_2) = w_2, \bar{\mu}(m_3) = m_3, \bar{\mu}(w_3) = w_3$ and $\bar{\mu}(m_1) >_{m_1} \mu(m_1), \bar{\mu}(w_1) >_{w_1} \mu(w_1)$.

Hence (a) and (b) hold. □ □

Propositions 4 and 5 combined yield

Corollary 1. *Let μ be a matching in the marriage market $(M, W; \mathbf{P}; \mathcal{E}')$ that is individually rational but unstable. Then there exist a pair $\{m, w\} \in \mathcal{E}'$ blocking μ and a stable matching $\bar{\mu}$ in the marriage market $(M, W; \mathbf{P}; \mathcal{E}')$ such that (i) $\bar{\mu}(m) >_m \mu(m)$ and $\bar{\mu}(w) \geq_w \mu(w)$ or (ii) $\bar{\mu}(m) \geq_m \mu(m)$ and $\bar{\mu}(w) >_w \mu(w)$.*

After focusing on potential gainers from less restrictive marriage markets or rematching, let us pay attention to potential losers. The key result is the following.

Proposition 6. *Let $(M, W; \mathbf{P}; \mathcal{E})$ and $(M, W; \mathbf{P}; \mathcal{E}')$ be two marriage markets with strict preferences and $\mathcal{E} \subseteq \mathcal{E}'$. Let μ be a stable matching in $(M, W; \mathbf{P}; \mathcal{E})$ and μ' be a stable matching in $(M, W; \mathbf{P}; \mathcal{E}')$. Then*

$$|\{i \in I : \mu'(i) <_i \mu(i)\}| \leq |\{i \in I : \mu'(i) >_i \mu(i)\}|.$$

Proof. Let $i \in I$ with $\mu'(i) <_i \mu(i)$. Then individual rationality of μ' implies $\mu(i) \neq i$. Set $j = \mu(i)$. Then $\mu(j) = i$. $\mu'(j) <_j \mu(j)$ would yield $\mu'(i) <_i j$ and $\mu'(j) <_j i$, contradicting the stability of μ' . Hence $\mu'(j) >_j \mu(j)$. We have shown that $\mu(\{i \in I : \mu'(i) <_i \mu(i)\}) \subseteq \{i \in I : \mu'(i) >_i \mu(i)\}$ holds for the bijection μ . The assertion follows. $\blacksquare$

Hence, when feasibility constraints for matches are relaxed, the number of losers cannot exceed the number of gainers if preferences are strict. We will apply this result to the integration of marriage markets in Section 5. Proposition 6 can also be applied in the special case $\mathcal{E} = \mathcal{E}'$ and it implies then that the number of winners equals the number of losers when we compare two stable matchings for a given marriage market.⁷

4 Adding an Edge

Here we study the effect of adding just one edge $\{m, w\}$ to an existing bigraph $\mathcal{E}$. To be precise, we consider a marriage market $(M, W; \mathbf{P}; \mathcal{E})$, $m \in M$, $w \in W$ and a bigraph $\mathcal{E}'$ such that $\{m, w\} \notin \mathcal{E}$ and $\mathcal{E}' = \mathcal{E} \cup \{\{m, w\}\}$.

Suppose first that μ' is a stable matching of the marriage market $(M, W; \mathbf{P}; \mathcal{E}')$. Then there are two possibilities. Either $\mu'(m) = w$ in which case μ' is not a matching of the marriage market $(M, W; \mathbf{P}; \mathcal{E})$. Or $\mu'(m) \neq w$ in which case μ' is a stable matching of the marriage market $(M, W; \mathbf{P}; \mathcal{E})$.

Suppose next that μ is a stable matching of the marriage market $(M, W; \mathbf{P}; \mathcal{E})$. Then there are again two possibilities. If $\{m, w\}$ does not block μ , then μ is a

⁷In the case $\mathcal{E} = \mathcal{E}'$, this result exists already in the literature; see Corollary 2.2.1 in Roth and Sotomayor (1990).

stable matching in the marriage market $(M, W; \mathbf{P}; \mathcal{E}')$. If $\{m, w\}$ blocks μ , then the following holds: First, μ is individually rational, since μ is stable in $(M, W; \mathbf{P}; \mathcal{E})$. Second, $\{m, w\}$ is the only element of $\mathcal{E}'$ that blocks μ . Hence by Corollary 1, there exists a stable matching $\bar{\mu}$ in the marriage market $(M, W; \mathbf{P}; \mathcal{E}')$ such that (i) $\bar{\mu}(m) >_m \mu(m)$ and $\bar{\mu}(w) \geq_w \mu(w)$ or (ii) $\bar{\mu}(m) \geq_m \mu(m)$ and $\bar{\mu}(w) >_w \mu(w)$. In case of strict preferences of m and w , (iii) $\bar{\mu}(m) >_m \mu(m)$ and $\bar{\mu}(w) >_w \mu(w)$. In general, other individuals may be affected by the reshuffling of matches when μ is replaced by $\bar{\mu}$. Let us examine several examples.

The first example is a variation of Example 2.19 in Roth and Sotomayor.

Example 4 (Effect of adding an edge). Let $k = 3$, $\ell = 4$, and preferences be given by the lists

$$P(m_1) = w_4, w_1, w_3; P(m_2) = w_3, w_2; P(m_3) = w_1, w_3;$$

$$P(w_1) = m_1, m_3; P(w_2) = m_2; P(w_3) = m_3, m_2; P(w_4) = m_2, m_1.$$

Further, let $\mathcal{E} = \mathcal{C} \setminus \{\{m_1, w_4\}\}$. Then

$$P^{\mathcal{E}}(m_1) = w_1, w_3; P^{\mathcal{E}}(m_2) = P(m_2); P^{\mathcal{E}}(m_3) = P(m_3);$$

$$P^{\mathcal{E}}(w_1) = P(w_1); P^{\mathcal{E}}(w_2) = P(w_2); P^{\mathcal{E}}(w_3) = P(w_3); P^{\mathcal{E}}(w_4) = m_2.$$

Then the marriage market $(M, W; \mathbf{P}^{\mathcal{E}})$ and, hence, $(M, W; \mathbf{P}; \mathcal{E})$ has the unique stable matching μ given by

$$\mu(w_1) = m_1, \mu(w_2) = m_2, \mu(w_3) = m_3, \mu(w_4) = w_4$$

whereas the marriage market $(M, W, \mathbf{P})$ has the unique stable matching $\bar{\mu}$ given by

$$\bar{\mu}(w_1) = m_3, \bar{\mu}(w_2) = w_2, \bar{\mu}(w_3) = m_2, \bar{\mu}(w_4) = m_1.$$

Individuals m_1 and w_4 , whose marriage was impossible under the restriction $\mathcal{E}$, are both better off without the restriction. So are m_2 and m_3 . All women except w_4 are worse off at the matching $\bar{\mu}$ than at μ . $\square \square$

Example 5. Suppose that $k = \ell$ and like in Example 1,

$$P(m) = w_1, \dots, w_k \text{ for all } m \in M,$$

$$P(w) = m_1, \dots, m_k \text{ for all } w \in W$$

Consider an edge $\{m_i, w_j\} \in \mathcal{C}$, $\mathcal{E} = \mathcal{C} \setminus \{\{m_i, w_j\}\}$, and $\mathcal{E}' = \mathcal{C}$. The assortative matching $\bar{\mu}$ given by $\bar{\mu}(m_s) = w_s$ for all $s \in \{1, \dots, k\}$ is the unique stable matching

in the marriage market $(M, W, \mathbf{P}) = (M, W, \mathbf{P}; \mathcal{E}')$. If $i \neq j$, then $\bar{\mu}$ is also the unique stable matching in the marriage market $(M, W, \mathbf{P}; \mathcal{E})$. If $i = j < k$, then the unique stable matching μ in the matching marriage market $(M, W, \mathbf{P}; \mathcal{E})$ is given by $\mu(m_i) = w_{i+1}$, $\mu(m_{i+1}) = w_i$ and $\mu(m_s) = w_s$ for $s \neq i, i+1$. Hence m_i and w_i are positively affected and m_{i+1} and w_{i+1} are negatively affected by the addition of the edge $\{m_i, w_j\}$ to the bigraph $\mathcal{E}$. If $i = j = k$, then the unique stable matching μ in the matching marriage market $(M, W, \mathbf{P}; \mathcal{E})$ is given by $\mu(m_s) = w_s$ for $s = 1, \dots, k-1$ while m_k and w_k remain unmatched. Hence in this case, the addition of the edge $\{m_k, w_k\}$ to the bigraph $\mathcal{E}$ benefits m_k and w_k and does not affect anybody else. $\square \square$

Example 6. Let $k = \ell \geq 3$ and preferences as in Example 2.

Put $\mathcal{E}' = \mathcal{C}$ and $\mathcal{E} = \mathcal{C} \setminus \{\{m, w\}\}$ for some $\{m, w\} \in \mathcal{C}$. (a) If $m = m_i, w = w_i$ for some $i \in \{1, \dots, k\}$, then μ_M and μ_W are the stable matchings in the marriage market $(M, W, \mathbf{P}) = (M, W, \mathbf{P}; \mathcal{E}')$ whereas μ_M is the only stable matching in $(M, W, \mathbf{P}; \mathcal{E})$. (b) If $m = m_i, w = w_{i+1}$ for some $i \in \{1, \dots, k\}$, then μ_M and μ_W are the stable matchings in the marriage market $(M, W, \mathbf{P}) = (M, W, \mathbf{P}; \mathcal{E}')$ whereas μ_W is the only stable matching in $(M, W, \mathbf{P}; \mathcal{E})$. (c) If neither (a) nor (b) applies, then μ_M and μ_W are the stable matchings in both the marriage market $(M, W, \mathbf{P}) = (M, W, \mathbf{P}; \mathcal{E}')$ and the marriage market $(M, W, \mathbf{P}; \mathcal{E})$. Hence adding an edge either affects nobody or affects everybody, depending on which stable matching materializes. $\square \square$

5 Segregation and Desegregation

Segregation of various sorts, be it racial, religious or political, restricts the marriage market. As an example for the latter, consider the case of South and North Korea or former West and East Germany. How would desegregation affect the marriage market? We are going to study this question in a simple example.

Conversely, choice of location and, hence, spatial segregation, may in part be driven by the anticipated matching opportunities. We will consider the possibility that some or all individuals are mobile whereas the rest of the population resides in fixed locations. We allow for choice of location as a first step followed by matching in the second step.

More specifically, we consider the case where the population I is segregated into two sub-populations $I_1 = M_1 \cup W_1$ and $I_2 = M_2 \cup W_2$, with $M_i = M \cap I_i$ and $W_i = W \cap I_i$. Marriages are only possible within I_1 and within I_2 . This restriction corresponds to the bigraph $\mathcal{E} = \mathcal{C}_1 \cup \mathcal{C}_2$ where $\mathcal{C}_i$ is the complete bigraph based on M_i and W_i .

5.1 Desegregation

Example 7 (Effect of desegregation). Let $k = \ell = 8$ and as in Example 1,

$$P(m) = w_1, w_2, \dots, w_k \text{ for all } m \in M;$$

$$P(w) = m_1, m_2, \dots, m_k \text{ for all } w \in W.$$

Suppose the population $I = M \cup W$ is segregated into two parts $I_1 = M_1 \cup W_1$ and $I_2 = M_2 \cup W_2$ where $M_1 = \{m_1, m_2, m_3, m_4\}$, $M_2 = \{m_5, m_6, m_7, m_8\}$, $W_1 = \{w_2, w_3, w_4, w_5\}$, $W_2 = \{w_1, w_6, w_7, w_8\}$. Marriages are only possible within I_1 and within I_2 . This restriction corresponds to the bigraph $\mathcal{E} = \mathcal{C}_1 \cup \mathcal{C}_2$ where $\mathcal{C}_1$ is the complete bigraph based on M_1 and W_1 and $\mathcal{C}_2$ is the complete bigraph based on M_2 and W_2 . Let $\mathbf{P}_1$ denote the restriction of the preference profile $\mathbf{P}^\mathcal{E}$ to members of I_1 and $\mathbf{P}_2$ denote the restriction of the preference profile $\mathbf{P}^\mathcal{E}$ to members of I_2 . The assumptions underlying Fact 1 hold for the marriage markets $(M_1, W_1; \mathbf{P}_1)$, $(M_2, W_2; \mathbf{P}_2)$ and $(M, W; \mathbf{P})$. Therefore:

Under segregation, there exists a unique stable matching $\mu^\mathcal{E}$ given by

$$\mu^\mathcal{E}(m_1) = w_2, \mu^\mathcal{E}(m_2) = w_3, \mu^\mathcal{E}(m_3) = w_4, \mu^\mathcal{E}(m_4) = w_5,$$

$$\mu^\mathcal{E}(m_5) = w_1, \mu^\mathcal{E}(m_6) = w_6, \mu^\mathcal{E}(m_7) = w_7, \mu^\mathcal{E}(m_8) = w_8.$$

After desegregation, there exists a unique stable matching μ given by

$$\mu(m_i) = w_i \text{ for } i = 1, \dots, 8.$$

Men m_1, m_2, m_3, m_4 and woman w_1 gain from desegregation whereas male m_5 and females w_2, w_3, w_4, w_5 lose. $\square \square$

In the foregoing example, one may consider limited mobility instead of desegregation. Suppose that I_1 and I_2 reside in far apart locations and know very little of each other. But the women in I_2 correctly assume that the men in I_1 are more attractive than the ones in I_2 and one member of W_2 decides to move to I_1 .

If w_1 joins I_1 , then at the new stable matching, still men m_1, m_2, m_3, m_4 and woman w_1 are better off than before and male m_5 and females w_2, w_3, w_4, w_5 are worse off than before, female w_5 now remaining single. But there are additional gainers and losers from w_1 's move. m_6, m_7 and m_8 are losers as well: m_6 is now matched with w_7 , m_7 is matched with w_8 , and m_8 now remains single. w_6, w_7 and w_8 are gainers. Hence everybody is affected by w_1 's move.

If w_8 instead of w_1 joins I_1 , then only w_8 and m_8 are affected, both remaining single. If rather w_7 moves, then w_7, m_7 and m_8 lose while w_8 gains. If only w_6 moves, then w_6, m_6, m_7 and m_8 lose while w_7 and w_8 gain.

5.2 Replica Marriage Markets

It is interesting to consider the special case where two sub-populations are identical in the sense that the merged marriage market is a two-fold replica of a basic marriage market $(M^0, W^0; \mathbf{P}^0)$. There are two identical copies $(M^A, W^A; \mathbf{P}^A)$ and $(M^B, W^B; \mathbf{P}^B)$ of $(M^0, W^0; \mathbf{P}^0)$ with pairwise disjoint sets M^A, W^A, M^B, W^B . For instance, if M^0 consists of males $m_1^0, \dots, m_k^0$, then M^A consists of corresponding males $m_1^A, \dots, m_k^A$. $m_i^A \in M^A$ strictly prefers $w_j^A \in W^A$ to $w_{j'}^A \in W^A$ if and only if m_i^0 strictly prefers w_j^0 to $w_{j'}^0$, etc. Again, we can analyze the two markets $(M^A, W^A; \mathbf{P}^A)$ and $(M^B, W^B; \mathbf{P}^B)$ separately and then compare the outcome with the outcome in the replica marriage market (common market) $(M^A \cup M^B, W^A \cup W^B; \mathbf{P})$ where, for example, for $G, H \in \{A, B\}$, m_i^A strictly prefers w_j^G to $w_{j'}^H$ if and only if m_i^0 strictly prefers w_j^0 to $w_{j'}^0$. In the common market, nobody has more attractive potential partners than in their respective separate market. Therefore, one might expect that stable matchings of the separate markets give rise to stable matchings in the common market. However, this is not necessarily the case.

Let μ^A be a stable matching in $(M^A, W^A; \mathbf{P}^A)$ and μ^B be a stable matching in $(M^B, W^B; \mathbf{P}^B)$. Define the matching $\mu = \mu^A \times \mu^B$ by $\mu(m^A) = \mu^A(m^A)$ for $m^A \in M^A$ and $\mu(m^B) = \mu^B(m^B)$ for $m^B \in M^B$. Then the question is whether $\mu^A \times \mu^B$ is stable. Obviously, it is if the below interchangeability condition (5.2d) holds and μ^A and μ^B are identical up to labeling, that is $\mu^A(m_i^A) = w_j^A$ if and only if $\mu^B(m_i^B) = w_j^B$ and $\mu^A(i^A) = i^A$ if and only if $\mu^B(i^B) = i^B$. $\mu^A \times \mu^B$ need not be stable in general. Within a given marriage market, the M -optimal stable matching

and the W -optimal stable matching are the ones that differ the most — in terms of male preferences or female preferences — among pairs of stable matchings, provided that they exist (and differ at all). Hence a combination of those two is possibly unstable. To examine this case, let us assume for the remainder of this subsection:

(5.2a) Preferences in the basic marriage market are strict.

(5.2b) Everybody finds all potential mates acceptable.

(5.2c) $|M^0| = |W^0| = k$.

(5.2d) *Interchangeability*: A person is indifferent between being matched with an A -partner and being matched with the corresponding B -partner.

Then all individuals are matched in a stable matching and the respective M -optimal stable matchings μ_M^A, μ_M^B and the respective W -optimal stable matchings μ_W^A, μ_W^B exist. The first two are identical up to labeling and the second two are identical up to labeling. Hence $\mu_M^A \times \mu_M^B$ and $\mu_W^A \times \mu_W^B$ are stable matchings.

Let us focus on $\mu_M^A \times \mu_W^B$ from now on. Then the following holds:

Proposition 7. *If the couple $\{m, w\}$ blocks $\mu_M^A \times \mu_W^B$, then $(m, w) \in M^B \times W^A$.*

Proof. We can rule out $(m, w) \in M^A \times W^A$ and $(m, w) \in M^B \times W^B$ because of the stability of μ^A and μ^B , respectively.

Suppose $(m, w) \in M^A \times W^B$. Let μ' be a stable matching in the common marriage market, which exists. Then by Proposition 4, $\mu'(m) >_m (\mu_M^A \times \mu_W^B)(m)$ or $\mu'(w) >_w (\mu_M^A \times \mu_W^B)(w)$. Without loss of generality, assume $\mu'(m) >_m (\mu_M^A \times \mu_W^B)(m) = \mu_M^A(m)$. (When $\mu'(w) >_w (\mu_M^A \times \mu_W^B)(w)$, we can proceed in an analogous way.)

CASE 1: $\mu'(m) \in W^A$.

Then we can renumber the individuals in $M^0 \cup W^0$ such that $\{1, \dots, k\}$ is partitioned into subsets $L_0 = \{1, \dots, \ell_0\}, \dots, L_n = \{\ell_{n-1} + 1, \dots, \ell_n = k\}$ (where L_0 or $L_1, \dots, L_n$ may be empty), $\mu'(\{m_i^A, m_i^B\}) = \{w_i^A, w_i^B\}$ for $i \in L_0$, $\mu'(\{m_i^A, m_i^B\}) = \{w_i^A, w_{i+1}^B\}$ for $\ell_{r-1} + 1 \leq i < \ell_r, 1 \leq r \leq n$, and $\mu'(\{m_{\ell_r}^A, m_{\ell_r}^B\}) = \{w_{\ell_r}^A, w_{\ell_{r-1}+1}^B\}$ for $1 \leq r \leq n$. Moreover, if $m \notin \{m_i^A | i \in L_0\}$ the renumbering can be such

that $m = m_{\ell_0+1}^A$ and $\mu'(m) = w_{\ell_0+1}^A$. Setting $\mu^A(m_i^A) = w_i^A$ defines a stable matching μ^A in $(M^A, W^A; \mathbf{P}^A)$, because of the stability of μ' in the common marriage market and the interchangeability condition. By construction, $\mu^A(m) = \mu'(m) >_m \mu_M^A(m)$. But this contradicts the fact that $\mu_M^A(m)$ is the best partner m gets in all stable matchings of $(M^A, W^A; \mathbf{P}^A)$.

CASE 2: $\mu'(m) \in W^B$.

Then we can renumber the individuals in $M^0 \cup W^0$ such that $\{1, \dots, k\}$ is partitioned into subsets $L_0 = \{1, \dots, \ell_0\}, \dots, L_n = \{\ell_{n-1} + 1, \dots, \ell_n = k\}$ (where L_0 or $L_1, \dots, L_n$ may be empty), $\mu'(\{m_i^A, m_i^B\}) = \{w_i^A, w_i^B\}$ for $i \in L_0$, $\mu'(\{m_i^A, m_i^B\}) = \{w_i^B, w_{i+1}^A\}$ for $\ell_{r-1} + 1 \leq i < \ell_r$, $1 \leq r \leq n$, and $\mu'(\{m_{\ell_r}^A, m_{\ell_r}^B\}) = \{w_{\ell_r}^B, w_{\ell_{r-1}+1}^A\}$ for $1 \leq r \leq n$. Moreover, if $m \notin \{m_i^A \mid i \in L_0\}$ the renumbering can be such that $m = m_{\ell_0+1}^A$ and $\mu'(m) = w_{\ell_0+1}^B$. Setting $\mu^A(m_i^A) = w_i^A$ defines a stable matching μ^A in $(M^A, W^A; \mathbf{P}^A)$, because of the stability of μ' in the common marriage market and the interchangeability condition. By construction (and the interchangeability condition), $\mu^A(m) = w_1^A \sim_m w_1^B = \mu'(m) >_m \mu_M^A(m)$. But this contradicts again the fact that $\mu_M^A(m)$ is the best partner m gets in all stable matchings of $(M^A, W^A; \mathbf{P}^A)$.

We arrive at a contradiction in both cases. Hence, to the contrary, $(m, w) \notin M^A \times W^B$. $(m, w) \in M^B \times W^A$ is the only remaining alternative. $\blacksquare$

Take for instance, similar to Example 2.19 in Roth and Sotomayor (1990), $k = 4$ and

$$\begin{array}{ll} P^0(m_1^0) = w_1^0, w_2^0, w_3^0, w_4^0 & P^0(w_1^0) = m_2^0, m_3^0, m_1^0, m_4^0 \\ P^0(m_2^0) = w_4^0, w_2^0, w_3^0, w_1^0 & P^0(w_2^0) = m_3^0, m_1^0, m_2^0, m_4^0 \\ P^0(m_3^0) = w_4^0, w_3^0, w_1^0, w_2^0 & P^0(w_3^0) = m_4^0, m_1^0, m_2^0, m_3^0 \\ P^0(m_4^0) = w_1^0, w_4^0, w_3^0, w_2^0 & P^0(w_4^0) = m_1^0, m_4^0, m_2^0, m_3^0 \end{array}$$

Then $\mu_M^A(w_i^A) = m_i^A$ for $i = 1, 2, 3, 4$; $\mu_W^B(m_1^B) = w_4^B$, $\mu_W^B(m_2^B) = w_1^B$, $\mu_W^B(m_3^B) = w_2^B$, $\mu_W^B(m_4^B) = w_3^B$, and the pair $\{m_2^B, w_3^A\}$ is blocking $\mu_M^A \times \mu_W^B$.

Next we present two cases in which $\mu_M^A \times \mu_W^B$ is stable. First, take Example 2 as the basic marriage market model. In that case, each man gets his most preferred partner and each woman gets her second most preferred partner in the M -optimal matching; each woman gets her most preferred partner and each man

gets his second most preferred partner in the W -optimal matching. If a pair $\{m, w\}$ blocks $\mu_M^A \times \mu_W^B$, then $(m, w) \in M^B \times W^A$ by Proposition 7. Moreover, $w >_m \mu_W^A(m)$, $m >_w \mu_M^B(w)$, and consequently, m and w are mutually most preferred partners. But then the corresponding pairs would belong to every stable matching in $(M^A, W^A; \mathbf{P}^A)$ and $(M^B, W^B; \mathbf{P}^B)$. In particular, they would belong to μ_M^A and μ_W^B . Hence to the contrary, $\{m, w\}$ does not block $\mu_M^A \times \mu_W^B$.

Second, take Example 2.10 of Roth and Sotomayor (1990) as the basic marriage market model: $k = 3$ and

$$\begin{aligned} P^0(m_1^0) &= w_1^0, w_2^0, w_3^0 & P^0(w_1^0) &= m_1^0, m_2^0, m_3^0 \\ P^0(m_2^0) &= w_1^0, w_2^0, w_3^0 & P^0(w_2^0) &= m_1^0, m_3^0, m_2^0 \\ P^0(m_3^0) &= w_1^0, w_3^0, w_2^0 & P^0(w_3^0) &= m_1^0, m_2^0, m_3^0 \end{aligned}$$

Then $\mu_M^A(w_i^A) = m_i^A$ for $i = 1, 2, 3$, $\mu_W^B(m_1^B) = w_1^B$, $\mu_W^B(m_2^B) = w_3^B$, $\mu_W^B(m_3^B) = w_2^B$, and $\mu_M^A \times \mu_W^B$ cannot be blocked.

The two cases share several features. Among other things, there exist only two stable matchings in the basic marriage market, the M -optimal stable matching μ_M^0 and the W -optimal stable matching μ_W^0 which in these cases are obtained via the following iterative procedure. In a first step, match all pairs consisting of mutually most preferred partners. If none exist, go to the last step. Else proceed to the second step and match all pairs consisting of mutually most preferred partners in the unmatched population. If none exist, go to the last step. Else proceed to the next step. The procedure continues this way until all individuals are matched (the remaining population is empty) or the final step is reached. In the former case, the unique stable matching has been determined. In the latter case, the unmatched population consists of least two men and two women. μ_M^0 results if every remaining man is matched with his best choice among the remaining women while every remaining woman is matched with her second choice among the remaining men. In μ_W^0 , every remaining woman is matched with her best choice among the remaining men while every remaining man is matched with his second choice among the remaining women. If μ_M^0 and μ_W^0 can be obtained in this way, then $\mu_M^A \times \mu_W^B$ is stable, by the argument given in the first case above.

5.3 More on Integration of Marriage Markets

In the preceding subsection, we have found that integration of marriage markets can destabilize local matchings when the integration of marriage markets concerns merged marriage markets with identical sub-populations, that is, when conditions (5.a) – (5.d) hold. In this subsection, we show that such destabilizing forces may be more or less pronounced when we consider an alternative set-up. Local markets are structurally identical but they are not identical copies: The interchangeability condition (5.d) fails to hold. We consider the following example.

Example 8. In this example we consider variations based on the local marriage market

$$\begin{aligned} P^0(m_1^0) &= w_1^0, w_2^0 \\ P^0(m_2^0) &= w_2^0, w_1^0 \\ P^0(w_1^0) &= m_2^0, m_1^0 \\ P^0(w_2^0) &= m_1^0, m_2^0 \end{aligned}$$

where the men-optimal stable matching μ_M^0 and the women-optimal stable matching μ_W^0 are given by

$$\begin{aligned} \mu_M^0(m_i^0) &= w_i^0 \quad \text{for } i = 1, 2 \\ \mu_W^0(w_1^0) &= m_2^0, \quad \mu_W^0(w_2^0) = m_1^0. \end{aligned}$$

This constitutes an instance where those two matchings are obtained via the construction outlined at the end of subsection 5.2 and, hence, $\mu_M^A \times \mu_W^B$ and $\mu_W^A \times \mu_M^B$ are stable matchings in the common marriage market in case the interchangeability condition (5.d) holds.

We now consider integration of two structurally equivalent local marriage markets, i.e., integration of two marriage markets with the same preferences when restricted to the local market. The local markets are again denoted by $(M^A, W^A; \mathbf{P}^A)$ and $(M^B, W^B; \mathbf{P}^B)$ and stable local matchings are denoted μ^A and μ^B .

CASE A: STABILITY

Consider the common marriage market

$$\begin{aligned}
P(m_1^A) &= w_1^A, w_2^A, w_1^B, w_2^B & P(w_1^A) &= m_2^A, m_1^A, m_2^B, m_1^B \\
P(m_2^A) &= w_2^A, w_1^A, w_2^B, w_1^B & P(w_2^A) &= m_1^A, m_2^A, m_1^B, m_2^B \\
P(m_1^B) &= w_1^B, w_2^B, w_1^A, w_2^A & P(w_1^B) &= m_2^B, m_1^B, m_2^A, m_1^A \\
P(m_2^B) &= w_2^B, w_1^B, w_2^A, w_1^A & P(w_2^B) &= m_1^B, m_2^B, m_1^A, m_2^A.
\end{aligned}$$

There are four possible stable matchings of the form $\mu^A \times \mu^B$ given by the cases where μ^A and μ^B are each either the men-optimal or women-optimal stable matchings in their respective local markets. These matchings remain stable after market integration. The same observations hold for the common marriage markets

$$\begin{aligned}
P(m_1^A) &= w_1^A, w_1^B, w_2^A, w_2^B & P(w_1^A) &= m_2^A, m_2^B, m_1^A, m_1^B \\
P(m_2^A) &= w_2^A, w_2^B, w_1^A, w_1^B & P(w_2^A) &= m_1^A, m_1^B, m_2^A, m_2^B \\
P(m_1^B) &= w_1^B, w_1^A, w_2^B, w_2^A & P(w_1^B) &= m_2^B, m_2^A, m_1^B, m_1^A \\
P(m_2^B) &= w_2^B, w_2^A, w_1^B, w_1^A & P(w_2^B) &= m_1^B, m_1^A, m_2^B, m_2^A
\end{aligned}$$

and

$$\begin{aligned}
P(m_1^A) &= w_1^A, w_1^B, w_2^B, w_2^A & P(w_1^A) &= m_2^A, m_2^B, m_1^B, m_1^A \\
P(m_2^A) &= w_2^A, w_2^B, w_1^B, w_1^A & P(w_2^A) &= m_1^A, m_1^B, m_2^B, m_2^A \\
P(m_1^B) &= w_1^B, w_1^A, w_2^A, w_2^B & P(w_1^B) &= m_2^B, m_2^A, m_1^A, m_1^B \\
P(m_2^B) &= w_2^B, w_2^A, w_1^A, w_1^B & P(w_2^B) &= m_1^B, m_1^A, m_2^A, m_2^B.
\end{aligned}$$

CASE B: INSTABILITY

Consider next the common marriage market

$$\begin{aligned}
P(m_1^A) &= w_1^B, w_2^B, w_1^A, w_2^A & P(w_1^A) &= m_2^B, m_1^B, m_2^A, m_1^A \\
P(m_2^A) &= w_2^B, w_1^B, w_2^A, w_1^A & P(w_2^A) &= m_1^B, m_2^B, m_1^A, m_2^A \\
P(m_1^B) &= w_1^A, w_2^A, w_1^B, w_2^B & P(w_1^B) &= m_2^A, m_1^A, m_2^B, m_1^B \\
P(m_2^B) &= w_2^A, w_1^A, w_2^B, w_1^B & P(w_2^B) &= m_1^A, m_2^A, m_1^B, m_2^B.
\end{aligned}$$

Now, none of the four matchings $\mu^A \times \mu^B$ from the previous case is stable after the integration of the marriage markets. Individuals exhibit a strict preference for foreign partners. Therefore, the couple $\{m_1^A, w_1^B\}$ blocks $\mu_M^A \times \mu_W^B$ and the

assertion of Proposition 7 does not hold in this case.⁸

CASE C: PARTIAL STABILITY

Consider the common marriage market

$$\begin{aligned} P(m_1^A) &= w_1^B, w_1^A, w_2^B, w_2^A & P(w_1^A) &= m_2^B, m_1^B, m_2^A, m_1^A \\ P(m_2^A) &= w_2^B, w_2^A, w_1^B, w_1^A & P(w_2^A) &= m_1^B, m_2^B, m_1^A, m_2^A \\ P(m_1^B) &= w_1^A, w_1^B, w_2^B, w_2^A & P(w_1^B) &= m_2^B, m_2^A, m_1^B, m_1^A \\ P(m_2^B) &= w_2^A, w_2^B, w_1^B, w_1^A & P(w_2^B) &= m_1^B, m_1^A, m_2^B, m_2^A. \end{aligned}$$

Here, of the four cases of the form $\mu^A \times \mu^B$ considered previously, three are stable after integration of marriage markets. However, $\mu_W^A \times \mu_M^B$ is unstable after market integration. It is blocked by the couple $\{m_2^A, w_1^B\}$. $\square \square$

5.4 Winners and Losers from Market Integration

We next examine winners and losers from integration of marriage markets. Again, as in subsection 5.2, previous stable matchings can only be destabilized through new feasible matches in the global marriage market. In Case B of the last subsection, all individuals gain from market integration. As a rule, however, there are gainers and losers. According to Proposition 6, if preferences are strict, then there cannot be more losers than gainers, which implies that not more than half of the population consists of losers. The latter conclusion can be strengthened without any assumption on preferences:

Proposition 8. *Suppose the population is segregated into $K \geq 2$ non-empty subpopulations $I_1, \dots, I_K$. Let μ be a matching under segregation and μ' be a stable matching after desegregation. Then*

$$|\{i \in I_k : \mu'(i) <_i \mu(i)\}| \leq |I_k|/2 \text{ for } k = 1, \dots, K.$$

Proof. Suppose that $|\{i \in I_k : \mu'(i) <_i \mu(i)\}| > |I_k|/2$ for some $k \in \{1, \dots, K\}$. $\mu'(i) <_i \mu(i)$ implies $\mu(i) \neq i$ for $i \in I_k$. That is, all members of $J_k \equiv \{i \in I_k : \mu'(i) <_i \mu(i)\}$ were matched under μ . Since $|J_k| > |I_k|/2$, two members of J_k were

⁸The proof of Proposition 7 relies on the interchangeability condition (5.d).

matched with each under μ . Therefore, they can block μ' , contradicting the stability of μ' . Hence to the contrary, the assertion of the proposition has to hold. ■

We thus observe a stark contrast between economic integration and the integration of marriage markets. Under standard assumptions, not all consumers can lose from the integration of originally separate pure exchange economies — which is a consequence of the first welfare theorem. However, almost all consumers may lose in some instances. While there can also be gainers and losers from the integration of marriage markets, there cannot be more losers than gainers if preferences are strict. Proposition 8 contains a stronger result as it states that the number of losers in each subpopulation cannot exceed the number of gainers in this subpopulation.

Integration of marriage markets may be more explicitly modeled as a two stage process in which local matchings occur in the first stage and global matchings in the second stage. Indeed the previous considerations can be put into such a two-stage process. Moreover, one might introduce plausible dependencies of matchings in the second stage on previous local matchings. An obvious candidate are adaptive preferences. Let us consider the previous scenario, but assume that while local markets are structurally identical, all individuals have adaptive preferences regarding matches in the following sense. A matched individual in a local market prefers to be single in the global market rather than to be matched with a partner they prefer less than their partner in the original local market.⁹ As a consequence, no individual will accept a match in the global market that would make him/her worse off compared to the situation in the local market. Adaptive preferences impose an extra condition for the admissibility of new stable matchings in the global market, namely the condition that no individual may be worse off in the global matching than he or she was in her respective status quo local matching.

In the previous situations of Example 8 adaptive preferences would not affect the conclusions regarding the stability or instability of local matchings. However, adaptive preferences can determine which type of global matchings arise and how they depend on the local matchings as we show in the next example.

⁹The notion of adaptive preferences goes back at least to von Weizsäcker (1971).

Example 9. The local marriage market is as in Example 8. Consider the common marriage market

$$\begin{aligned}
P(m_1^A) &= w_1^A, w_1^B, w_2^A, w_2^B & P(w_1^A) &= m_1^B, m_2^B, m_2^A, m_1^A \\
P(m_2^A) &= w_2^A, w_1^A, w_1^B, w_2^B & P(w_2^A) &= m_1^A, m_2^A, m_1^B, m_2^B \\
P(m_1^B) &= w_1^A, w_1^B, w_2^B, w_2^A & P(w_1^B) &= m_1^A, m_2^B, m_1^B, m_2^A \\
P(m_2^B) &= w_2^B, w_1^B, w_1^A, w_2^A & P(w_2^B) &= m_1^B, m_2^B, m_1^A, m_2^A.
\end{aligned}$$

Any stable matching in the common market will involve the pair (m_1^B, w_1^A) . The other elements of a stable matching¹⁰ depend on the initial local matchings. Consider, e.g., the men-optimal stable matchings. Then, the only stable matching in the common market is $(m_1^A), (m_2^A, w_2^A), (m_1^B, w_1^A), (m_2^B, w_2^B), (w_1^B)$. Since m_1^A cannot match with w_1^A anymore, he prefers to remain single causing w_1^B also to be single despite the fact that w_1^B was the second best choice of m_1^A and m_1^B the best choice of w_2^B . Such phenomena could not occur with standard preferences where everybody would be matched. $\square\square$

¹⁰In general, existence of stable matchings in the common market is guaranteed. Local matchings affect the preferences in the common market. However, once these preferences are determined, the standard existence result applies.

Part II

Endogenous Feasibility Constraints

6 Sabotaging Possible Matches

6.1 Definitions

As already discussed in the introduction, individuals have various ways to make particular matches difficult or infeasible. In this section, we investigate the consequences when individuals can sabotage a potential match, i.e., render the particular match infeasible. We consider the following scenario. Suppose a marriage market $\Gamma = (M, W, \mathbf{P})$ is given and individuals anticipate that a particular matching might realize. Then each woman decides whether to sabotage a specific edge (match). This gives rise to a restricted marriage market. We can embody the restrictions by way of preferences so that the new market becomes $\Gamma^S = (M, W, \mathbf{P}^S)$, where $\mathbf{P}^S$ arises from $\mathbf{P}$ by the elimination of the individuals from a person's preference list with whom the person can no longer be matched. We explore which stable matchings occur in the new market Γ^S when sabotage activities are optimally chosen. We focus on sabotage by one side of the market. We start by the following definition.

Definition 4. A **sabotage** (of women) is a mapping $S : W \rightarrow E(\Gamma) \cup \{\emptyset\}$, where $S(w) = \{m', w'\}$ means that w chooses to sabotage the edge $\{m', w'\}$ and $S(w) = \emptyset$ means that w takes no action. The women w for whom $S(w) \neq \emptyset$ will be called **saboteurs** and $S(w)$ is called the **sabotage strategy** of w .

We note that two women may sabotage the same match, which renders one of their actions redundant. A woman may sabotage matches to which she belongs. That is, $w \in S(w)$ is allowed. We next define best responses regarding sabotage strategies. For this purpose, we introduce the following notation.

For a sabotage S , let $\mathcal{S}^S$ denote the stable matchings in $(M, W, \mathbf{P}^S)$. For a sabotage S , $w' \in W$ and $S'(w') \in E(\Gamma) \cup \{\emptyset\}$, let $S|S'(w')$ denote the sabotage where all women $w \neq w'$ choose $S(w)$ and w' chooses $S'(w')$.

We can now define a best response.

Definition 5. Given a sabotage S and $\mu_S \in \mathcal{S}^S$, woman w' plays a **best response** if $\nexists S'(w') \in E(\Gamma) \cup \{\emptyset\}$, $\mu' \in \mathcal{S}^{S'(w')}$ such that $S'(w') \neq S(w')$ and $\mu'(w') >_{w'} \mu_S(w')$.

We next introduce three types of sabotage that involve stable matchings which are immune against individual incentives to sabotage.

Definition 6.

- A sabotage S is called a **Nash sabotage with respect to** $\mu_S \in \mathcal{S}^S$, if every woman plays a best response given S and μ_S .
- A sabotage S is a **Nash sabotage** if it is a Nash sabotage with respect to some $\mu_S \in \mathcal{S}^S$.
- A sabotage S is a **global Nash sabotage** if it is a Nash sabotage with respect to all $\mu_S \in \mathcal{S}^S$.

If $\mu_S \in \mathcal{S}^S$ and S is a Nash sabotage with respect to μ_S , we call μ_S immune against further sabotage. If in addition $S(w) = \emptyset$ for all $w \in W$, we call μ_S immune against (individual) sabotage.

6.2 Examples

We explore the consequences of sabotage with a few illustrative examples.

Example 10. Consider the following marriage market:

$$\begin{aligned} P(m_1) &= w_1, w_2, w_3 & P(w_1) &= m_2, m_1, m_3 \\ P(m_2) &= w_2, w_3, w_1 & P(w_2) &= m_1, m_2, m_3 \\ P(m_3) &= w_2, w_3, w_1 & P(w_3) &= m_2, m_3, m_1. \end{aligned}$$

The men-optimal stable matching is $\mu(m_i) = w_i$ for $i = 1, 2, 3$. Let us consider whether this matching is immune against sabotage. We observe that w_2 can strictly benefit from sabotaging the match $\{m_1, w_1\}$. Indeed, let us consider the sabotage $S(w_1) = \emptyset$, $S(w_2) = \{m_1, w_1\}$, $S(w_3) = \emptyset$. Then, there exists a unique stable matching in the post-sabotage market: $\mu(m_1) = w_2$, $\mu(m_2) = w_3$ and $\mu(m_3) = w_1$.

which strictly benefits w_2 . Hence, the men-optimal stable matching is not immune against sabotage.

We next consider stable matchings that are immune to further sabotage. Let $S(w_1) = \{m_2, w_2\}$, $S(w_2) = \{m_2, w_3\}$, $S(w_3) = \emptyset$ be a sabotage which yields the post-sabotage market

$$\begin{aligned} P(m_1) &= w_1, w_2, w_3 & P(w_1) &= m_2, m_1, m_3 \\ P(m_2) &= w_1 & P(w_2) &= m_1, m_3 \\ P(m_3) &= w_2, w_3, w_1 & P(w_3) &= m_3, m_1. \end{aligned}$$

In this market, there is one stable matching μ_S : $\mu_S(m_1) = w_2$, $\mu_S(m_2) = w_1$, $\mu_S(m_3) = w_3$. Both active saboteurs obtain their best possible match and thus the sabotage strategies are best responses. Given that the match $\{m_2, w_3\}$ is sabotaged, w_3 cannot improve its match $\{m_3, w_3\}$ by sabotaging another not yet sabotaged match. Hence, S is a Nash sabotage with respect to μ_S . Moreover, since μ_S is unique, S is a global Nash sabotage.

□ □

Example 11. Consider the following marriage market.

$$\begin{aligned} P(m_1) &= w_2, w_1, w_3 & P(w_1) &= m_1, m_2, m_3 \\ P(m_2) &= w_1, w_3, w_2 & P(w_2) &= m_2, m_1, m_3 \\ P(m_3) &= w_1, w_2, w_3 & P(w_3) &= m_1, m_2, m_3. \end{aligned}$$

Consider the stable matching $\mu(m_1) = w_2$, $\mu(m_2) = w_1$ and $\mu(m_3) = w_3$ which is the women- and men-optimal stable matching. Renouncing sabotaging is not a Nash sabotage with respect to this matching. Indeed, it would be a profitable deviation for w_3 to sabotage $\{m_1, w_2\}$ or $\{m_2, w_1\}$. Thus, even if women have their best possible stable matching in a given market, renouncing sabotaging by women is not a Nash sabotage.

Consider the sabotage strategies $S(w_1) = \emptyset$, $S(w_2) = \{m_2, w_1\}$ and $S(w_3) = \{m_1, w_1\}$. Then, there is a post-sabotage stable matching $\mu_S(w_1) = m_3$, $\mu_S(w_2) = m_2$, $\mu_S(w_3) = m_1$ that is strictly profitable for saboteurs w_2 and w_3 compared to the sabotage-free stable matching. On the other hand, w_1 is worse off than in the stable matching when no sabotage takes place. But she cannot do anything to prevent this. She could also sabotage some matches, but she would not benefit.

We note that there is another stable post-sabotage matching, namely $\mu'(w_1) = m_3$, $\mu'(w_2) = m_1$, $\mu'(w_3) = m_2$ which is strictly beneficial only for w_3 compared to the stable matching without sabotage. However, the above Nash sabotage strategies are also best responses with regard to the second possible stable post-sabotage matching. Hence, $S(w_1) = \emptyset$, $S(w_2) = \{m_2, w_1\}$ and $S(w_3) = \{m_2, w_1\}$ is a global Nash sabotage. $\square\square$

Example 12. Consider the following marriage market.

$$\begin{aligned} P(m_1) &= w_1, w_2, w_3 & P(w_1) &= m_2, m_1, m_3 \\ P(m_2) &= w_2, w_1, w_3 & P(w_2) &= m_1, m_2, m_3 \\ P(m_3) &= w_1, w_2, w_3 & P(w_3) &= m_2, m_1, m_3. \end{aligned}$$

The men-optimal stable matching is given by $\mu(m_i) = w_i$ for $i = 1, 2, 3$ and the women-optimal stable matching is $\mu(w_1) = m_2$, $\mu(w_2) = m_1$ and $\mu(w_3) = m_3$.

Let us consider the men-optimal stable matching. Individual w_1 or w_2 has strict incentives to sabotage by $S(w_1) = \{w_2, m_2\}$, and $S(w_2) = \{w_1, m_1\}$, respectively, to induce their optimal match. Hence, the men-optimal match is not immune against sabotage. However, $S(w_1) = \{w_2, m_2\}$ and $S(w_2) = \{w_1, m_2\}$ together cannot be part of a Nash sabotage. With $S(w_3) = \{m_2, w_1\}$ individual w_3 can ensure the post sabotage matching $\{m_1, w_2\}$, $\{m_2, w_3\}$, $\{m_3, w_1\}$ which would be beneficial for w_3 but is not a best response for w_1 .

There is a Nash sabotage S' where w_2 and w_3 act in a complementary way: $S'(w_2) = \{m_1, w_1\}$, $S'(w_3) = \{m_2, w_1\}$, $S'(w_1) = \emptyset$. This forces w_1 to match with her third choice m_3 in the resulting market. The unique post-sabotage stable matching is $\mu_S(m_1) = w_2$, $\mu_S(m_2) = w_3$ and $\mu_S(m_3) = w_1$. Hence, w_3 and w_2 will both profit from the sabotage and w_1 cannot avoid being harmed by sabotage of the other two women.

Consider now the women-optimal stable matching. Then, this matching is immune to further sabotage with $S(w_1) = \{m_1, w_3\}$, $S(w_2) = \{m_2, w_3\}$ and $S(w_3) = \emptyset$ being a Nash sabotage. In this situation, the two women who have their first choice defend their advantageous position in the marriage market through sabotage as they force w_3 to remain with her least preferred choice. Because each of the saboteurs keeps her best choice and w_3 cannot improve her situation, the above strategies constitute indeed a Nash sabotage. $\square\square$

Example 13. Consider the following marriage market:

$$\begin{aligned} P(m_1) &= w_1, w_2, w_3 & P(w_1) &= m_3, m_2, m_1 \\ P(m_2) &= w_2, w_3, w_1 & P(w_2) &= m_1, m_3, m_2 \\ P(m_3) &= w_3, w_1, w_2 & P(w_3) &= m_2, m_1, m_3. \end{aligned}$$

The men-optimal stable matching is $\{m_1, w_1\}, \{m_2, w_2\}, \{m_3, w_3\}$ while the women-optimal stable matching is $\{m_1, w_2\}, \{m_2, w_3\}, \{m_3, w_1\}$. Consider the men-optimal stable matching. Then, each woman has an incentive to sabotage its own match. Consider, e.g., $S(w_1) = \{m_1, w_1\}$. Then, the women-optimal match is the unique stable match in the post-sabotage market – and all three women strictly benefit. Indeed, the sabotage in which one woman destroys the match she would end up with in the men-optimal matching yields the women-optimal matching and is a global Nash sabotage.¹¹

Moreover, the complete sabotage $S(w_1) = \{m_1, w_1\}$, $S(w_2) = \{m_2, w_2\}$ and $S(w_3) = \{m_3, w_3\}$ of all matches in the men-optimal stable matching also yields the women-optimal stable matching and this matching is unique in the post-sabotage market. Hence, S is a global Nash sabotage. $\square\square$

6.3 General Considerations

The preceding examples indicate that Nash sabotage can take different forms. They show that under a Nash sabotage, a woman may be worse off than under no sabotage; that a women-optimal stable matching need not be immune against sabotage; that a woman can benefit from sabotaging matches she is involved in. Obviously, some men may benefit from sabotage by women. Since sabotage restricts matching opportunities, Proposition 6 implies that there cannot be more gainers than losers from sabotage.

Several open questions remain:

¹¹While we focus on a static analysis of sabotage, these considerations and the other examples could also be embedded in a dynamic setting. Suppose that in a first-stage, stable matchings are realized in which sabotage is not possible. Then, in a second stage, women decide on sabotage activities and a stable matching in the post-sabotage market occurs. If in this set-up, the men-optimal stable matching is realized in stage 1, some or all women would sabotage their own match and the women-optimal stable matching would occur in stage 2.

- Does a Nash sabotage always exist?
- When does no woman lose under a Nash sabotage compared to stable matchings without sabotage?
- Can a group of women benefit from coordinating their sabotage strategies?

Two partial results can be established:

Lemma 1. *Suppose that $\mu'_M (\mu'_W)$ is the status quo matching in the marriage market (M, W, P') . Moreover, suppose that $P \leq_W P'$, and that P arises from P' via female sabotage. If $\mu_M (\mu_W)$ is the matching resulting after sabotage, then the women do not lose.*

In this Lemma, $P \leq_W P'$ means that P is a sublist of the preference list P' by deleting entries only from the end of $P(w)$ for all $w \in W$. The Lemma follows from the combination of Theorem 2.24 and the Decomposition Lemma 2.20 in Roth and Sotomayor (1990) applied to W . We note – as shown in Example 11 – that it can not be excluded in general that some woman may lose through sabotaging even if we start from the men-optimal stable matching.

The next property shows that a coalition of women could benefit by coordinating their sabotage strategies.

Lemma 2. *Suppose a woman w_0 has at most $n-1$ acceptable men in her preference list in a marriage market (M, W, P') . Let (M, W, P) denote the marriage market resulting from the deletion of the complete preference list of w_0 via coordinated sabotage from the other $n-1$ women. Then $\mu_W \geq_w \mu'_W$ and $\mu_M \geq_w \mu'_M$ under P for every $w \neq w_0$.*

The preceding Lemma is a consequence of Theorem 2.25 in Roth and Sotomayor (1990), since the coordinated sabotage essentially eliminates a certain woman's preference list from the marriage market.

7 Segregation: Location Choice and Matching

We next investigate a second force that endogenously determines feasibility constraints on matching. That is, individuals decide in a first stage to which subpopulation they want to belong. In a second stage matching occurs within the

subpopulation. Let the labels 1 and 2 stand for two endogenously formed groups or for the endogenous populations of two locations (territories). I_1 is the player population that belongs to group 1 (lives in territory 1) and I_2 is the player population that belongs to group 2 (lives in territory 2).

There is a population I , partitioned into a finite non-empty set M consisting of men and a finite non-empty set W consisting of women. We begin with the special case where everybody is mobile.

In the first stage, each individual i makes a binary choice $s_i \in S_i = \{1, 2\}$. An action profile $s = (s_i)_{i \in I} \in S = \prod_{i \in I} S_i$ determines $I_1(s) = \{i \in I : s_i = 1\}$, $I_2(s) = \{i \in I : s_i = 2\}$, $M_1(s) = \{i \in M : s_i = 1\}$, etc. so that I is partitioned into $I_1(s)$ and $I_2(s)$, M is partitioned into $M_1(s)$ and $M_2(s)$, W is partitioned into $W_1(s)$ and $W_2(s)$.

In the second stage, matching takes place. Marriages are only possible within $I_1(s)$ and within $I_2(s)$. This restriction corresponds to the bigraph $\mathcal{E}(s) = \mathcal{C}_1(s) \cup \mathcal{C}_2(s)$ where $\mathcal{C}_i(s)$ is the complete bigraph based on $M_i(s)$ and $W_i(s)$.

Definition 7. A pair (s, μ) is a **Stable Location-Partner Choice (SLPC)** under universal mobility if

- (i) $s \in S$;
- (ii) μ is a stable matching in the marriage market $(M, W, \mathbf{P}; \mathcal{E}(s))$;
- (iii) there is no $i \in I$ so that μ is unstable in the marriage market $(M, W, \mathbf{P}; \mathcal{E}(3 - s_i, s_{-i}))$.

Condition (iii) means that μ is not destabilized when one person switches locations. Condition (ii) implies that μ is individually rational. Consequently, μ is destabilized by i switching location if and only if there exists a person j such that $s_i \neq s_j$ and the couple $\{i, j\}$ blocks μ in the marriage market $(M, W, \mathbf{P}; \mathcal{E}(3 - s_i, s_{-i}))$. But given (ii), the latter is equivalent to $\{i, j\}$ blocking μ in the marriage market $(M, W, \mathbf{P})$. Hence (ii) and (iii) imply that μ is a stable matching in the unrestricted marriage market $(M, W, \mathbf{P})$. The question arises whether an SLPC rules out any stable matching μ . The answer is no. For take any stable matching μ in the unrestricted marriage market $(M, W, \mathbf{P})$. Let M_1 be any subset of M . Set

$s_i = 1$ if $i \in M_1 \cup \mu(M_1)$ and $s_i = 2$ otherwise. Then (s, μ) is an SLPC. Notice, however, that s and μ have to be consistent in an SLPC (s, μ) : $s_i = s_{\mu(i)}$ for $i \in I$.

Next we consider the general case where some individuals may be immobile. There are two exogenous disjoint subsets I_1^o and I_2^o of I such that $s_i = 1$ for $i \in I_1^o$ and $s_i = 2$ for $i \in I_2^o$. Let $I^o = I_1^o \cup I_2^o$, the set of immobile individuals. We still define $S_i = \{1, 2\}$ for $i \in I$. But now the individuals in I^o have no choice whereas each $i \in I \setminus I^o$ chooses $s_i \in S_i$ in the first stage. We modify the definition of an SLPC accordingly where S and $I_1(s), I_2(s), \dots, \mathcal{E}(s)$ for $s \in S$ are defined as before.

Definition 8. A pair (s, μ) is a **Stable Location-Partner Choice (SLPC)** if

- (i) $s \in S$, $s_i = 1$ for $i \in I_1^o$, $s_i = 2$ for $i \in I_2^o$;
- (ii) μ is a stable matching in the marriage market $(M, W, \mathbf{P}; \mathcal{E}(s))$;
- (iii) there is no $i \in I \setminus I^o$ so that μ is unstable in the marriage market $(M, W, \mathbf{P}; \mathcal{E}(3-s_i, s_{-i}))$.

Mutatis mutandis, the analysis of the special case $I^o = \emptyset$ applies to the general case. Define $M_1^o = M \cap I_1^o$, $M_2^o = M \cap I_2^o$, $W_1^o = W \cap I_1^o$, $W_2^o = W \cap I_2^o$, $\mathcal{E}^* = \{\{m, w\} : m \in M_1^o, w \in W_2^o\} \cup \{\{m, w\} : m \in M_2^o, w \in W_1^o\}$, and $\mathcal{E}^o = \mathcal{C} \setminus \mathcal{E}^*$. $\mathcal{E}^*$ is the set of matches that are a priori impossible given I_1^o and I_2^o while $\mathcal{E}^o$ is the set of matches that are a priori possible.

Proposition 9. A pair (s, μ) is an SLPC if and only if

- (i) $s \in S$, $s_i = 1$ for $i \in I_1^o$, $s_i = 2$ for $i \in I_2^o$;
- (iv) μ is a stable matching of the marriage market $(M, W, \mathbf{P}; \mathcal{E}^o)$;
- (v) $s_i = s_{\mu(i)}$ for $i \in I$.

Proof. Suppose (s, μ) satisfies (i) – (iii). By (ii), there does not exist a couple $\{m, w\}$ that satisfies $m \in M$, $w \in W$, $s_m = s_w$ and blocks μ . By (iii), there does not exist a couple $\{m, w\}$ that satisfies $m \in M$, $w \in W$, $s_m \neq s_w$, $\{m, w\} \notin \mathcal{E}^*$ and blocks μ . Hence there does not exist a couple $\{m, w\} \in \mathcal{E}^o$ that blocks μ . By (ii), μ is individually rational. Moreover, (i) implies $\mathcal{E}(s) \subseteq \mathcal{E}^o$. Therefore, (iv) holds. Finally, (ii) implies (v).

Conversely, suppose that (s, μ) satisfies (i), (iv) and (v). Since (i) implies $\mathcal{E}(s) \subseteq \mathcal{E}^o$, (i) and (iv) imply that μ is not blocked by any $\{m, w\} \in \mathcal{E}(s)$. By (v), every couple $\{m, w\}$ that is matched in μ satisfies $s_m = s_{\mu(m)} = s_w$, hence $\{m, w\} \in \mathcal{E}(s)$. By (iv), μ is individually rational. These properties combined render (ii). (iv) further implies (iii). ■

The proposition shows that the only restriction on matchings that belong to an SLPC is $\mathcal{E}^o$. Immobile persons can only be matched if they reside in the same location. Two mobile partners can always be matched if they want to and choose the same location. A mobile and an immobile partner can also be matched, provided the mobile individual chooses the partner's location.

As a corollary, the proposition yields existence of an SLPC. Namely, by Proposition 2, there exists a stable matching of the marriage market $(M, W, \mathbf{P}; \mathcal{E}^o)$. Set $s_i = 1$ for $i \in I_1^o \cup \mu(I_1^o)$ and $s_i = 2$ otherwise. Then (s, μ) constitutes an SLPC.

One can view the status quo in Example 7 as a situation where everybody is immobile. If a particular person, say w_1 becomes mobile, then all individuals are (positively or adversely) affected. Hence a change of mobility of just one person can have a drastic effect.

At additional notational costs, the foregoing analysis can be extended to models with more than two locations, with identical qualitative conclusions.

8 Preferences for Partners and Locations

In the previous section, we considered location choice followed by matching. The sequential treatment facilitates the exposition. Yet the formal analysis of that model does not depend on the sequencing of events. All that is required in an SLPC is the consistency condition (v): $s_i = s_{\mu(i)}$ for $i \in I$. Hence we may assume just as well that location choice and matching occur simultaneously.

The only role of location choice in Section 7 is to restrict the set of feasible matches to the bigraph $\mathcal{E}(s)$. Individuals do not care about their location otherwise. In reality, many people do care about their location, because of regional attachment, professional, cultural, recreational and social opportunities. That is, they have preferences for partners and locations. In this section therefore, we consider matching when individuals have preferences for partners and locations but

can only be matched to partners in the same location.

8.1 Definitions

For the time being we assume that two individuals have to reside in the same location in order to be matched. Let $\mathcal{X}$ denote a non-empty finite set of locations labeled $x_1, \dots, x_n$. There is again a finite population I that is partitioned into a non-empty male subpopulation $M = \{m_1, \dots, m_k\}$ and a non-empty female subpopulation $W = \{w_1, \dots, w_\ell\}$. Individuals have preferences for partners and locations. Each $m \in M$ has complete and transitive preferences on $(W \times \mathcal{X}) \cup \{m\}$, represented by an ordered list $P(m)$ as before. For instance, the abbreviated list

$$P(m) = (w_1, x_2), (w_2, x_1)$$

means that m 's first choice is to be married to woman w_1 at location x_2 ; his second choice is to be married to woman w_2 at location x_1 ; his third choice is to remain single.

Similarly, each $w \in W$ has complete and transitive preferences on $(M \times \mathcal{X}) \cup \{w\}$, represented by an ordered list $P(w)$. If an individual remains single, the person still has preferences for locations. But we need not list those preferences explicitly, assuming implicitly that the person will reside at one of her most preferred locations when single. We now say that $(m, x) \in M \times \mathcal{X}$ is acceptable to $w \in W$ if $(m, x) \geq_w w$ and that $(w, y) \in W \times \mathcal{X}$ is acceptable to $m \in M$ if $(w, y) \geq_m m$.

For $\mathcal{Z} \subseteq \mathcal{C} \times \mathcal{X}$, define $\mathcal{M}(\mathcal{Z}) = \{\{m, w\} \in \mathcal{C} \mid \exists x \in \mathcal{X} : (\{m, w\}, x) \in \mathcal{Z}\}$, the projection of $\mathcal{Z}$ to $\mathcal{C}$.

Definition 9. $\mathcal{Z} \subseteq \mathcal{C} \times \mathcal{X}$ is a **matching of persons and locations** if

(M1) $\mathcal{M}(\mathcal{Z})$ is a matching of men and women;

(M2) $(\{m, w\}, x) \in \mathcal{Z}$ implies that
 (m, x) is acceptable to w and
 (w, x) is acceptable to m .

The personal part of a matching, $\mathcal{M}(\mathcal{Z})$ has again an alternative representation by means of a bijection μ from I to I . Further, for a matching $\mathcal{Z}$ and $(\{m, w\}, x) \in \mathcal{Z}$ define $\lambda(m) = x$ and $\lambda(w) = x$. For an individual, $\lambda(i)$ stands for an unspecified

best location for i as a single person. Next we extend the concept of a stable matching.

Definition 10. *A matching $\mathcal{Z}$ is stable if:*

- (M3) *There is no matched individual with $i >_i (\mu(i), \lambda(i))$.*
- (M4) *There is no triple $(m, w, x) \in M \times W \times \mathcal{X}$ such that $(m, x) >_w (\mu(w), \lambda(w))$ and $(w, x) >_m (\mu(m), \lambda(m))$.*

If there is a unique location or if location choice is irrelevant, the concepts of matching and stable matching coincide with the classical ones of Gale and Shapley (1962).

8.2 Existence and Core Equivalence

We show in this subsection that much of the earlier analysis can be extended to the more general case. We demonstrate this for the two most important results, the existence of stable matchings and the characterization of stable matchings as core allocations.

Proposition 10 (Existence). *There exists a stable matching of persons and locations.*

Proof. We apply an adaptation of the deferred acceptance algorithm with men proposing. For each individual let us fix an order for the listing of alternatives between which the individual is indifferent. In the beginning, each man m with acceptable alternatives makes the proposal (m, x) to woman w if (w, x) is the first alternative on the man's preference list. Every woman rejects all proposals that are not acceptable to her and each woman who receives more than one acceptable proposal, rejects all but the one that comes first on her list. Any man whose proposal (m, x) is not rejected at this time is “engaged” at these terms.

Any man m whose proposal was rejected in the previous step proposes (m, x') to woman w' if (w', x') is first on his list of acceptable alternatives that did not cause a rejection yet. (If at any step, a man has exhausted his list of acceptable alternatives, he does not issue any more proposals.) Again, every woman rejects all new proposals that are not acceptable to her. Among the new acceptable proposals

plus (if applicable) the proposal underlying her engagement in the previous step, she rejects all but the one that comes first on her list. Any man whose proposal (m, x) is not rejected at this time is now “engaged” at these terms.

The algorithm stops after any step in which no proposal is rejected. Then every man is either engaged or all his proposals based on acceptable alternatives on his list have been rejected (including the case where there are no acceptable alternatives). All engaged couples now get married at the terms of the underlying proposals whereas all other individuals remain single.

The algorithm will eventually stop, since there are only a finite number of men, women, and locations and no man makes the same proposal more than once to any woman. The outcome is a matching. Let the matching be described by the mappings μ and λ . The matching is individually rational in the sense of condition (M3), since if m and w end up engaged, then m ’s underlying proposal (m, x) is acceptable to w and (w, x) is acceptable to m , hence $(m, x) \geq_w w$ and $(w, x) \geq_m m$. The matching produced by the algorithm also satisfies condition (M4). For suppose there exists $(m, w, x) \in M \times W \times \mathcal{X}$ with $(m, x) >_w (\mu(w), \lambda(w))$ and $(w, x) >_m (\mu(m), \lambda(m))$. Then (w, x) precedes $(\mu(m), \lambda(m))$ on m ’s preference list. Hence m proposed (m, x) to w prior to proposing $(m, \lambda(m))$ to $\mu(m)$. Since $(m, x) >_w (\mu(w), \lambda(w))$ and $(\mu(w), \lambda(w))$ is acceptable to w , (m, x) is acceptable to w — and preferred to $(\mu(w), \lambda(w))$. Hence from some step onward, w had engagements she preferred to the current match. Hence $(\mu(w), \lambda(w))$ cannot stipulate the terms of her final engagement. Therefore, $(m, x) >_w (\mu(w), \lambda(w))$ and $(w, x) >_m (\mu(m), \lambda(m))$ cannot both hold. This demonstrates (M4). ■

A standard result on two-sided matching is weak core equivalence. That is, the set of stable matchings and the weak core coincide. We can establish a weak core equivalence result for matching of persons and locations.

Let A denote a coalition, that is a non-empty subset of I . Consider two matchings $\mathcal{Z}$ and $\mathcal{Z}'$, described by pairs of mappings (μ, λ) and (μ', λ') , respectively. We say that the matching $\mathcal{Z}'$ dominates the matching $\mathcal{Z}$ if there exists a coalition A such that

$$\begin{aligned} \mu'(i) &\in A \text{ for all } i \in A; \\ (\mu'(i), \lambda'(i)) &>_i (\mu(i), \lambda(i)) \text{ for all } i \in A. \end{aligned}$$

In that case we say that μ' dominates μ on A . The weak core is the set of undom-

inated matchings.

Proposition 11 (Weak Core Equivalence).

The set of stable matchings and the weak core coincide.

Proof. Let $\mathcal{Z}$ be a matching that is unstable and described by (μ, λ) . Then (M3) is violated or (M4) is violated. In case (M3) is violated, there exists $i \in I$ with $i >_i (\mu(i), \lambda(i))$. Set $\mu'(j) = j$ for all $j \in I$. Then (μ', λ') dominates (μ, λ) on $\{i\}$. If (M4) is violated, there exists $(m, w, x) \in M \times W \times \mathcal{X}$ such that $(m, x) >_w (\mu(w), \lambda(w))$ and $(w, x) >_m (\mu(m), \lambda(m))$. Set $\mu'(m) = w, \mu'(w) = m, \lambda'(m) = \lambda'(w) = x$ and $\mu'(j) = j$ for $j \neq m, w$. Then (μ', λ') dominates (μ, λ) on $\{m, w\}$. Hence in any case, $\mathcal{Z}$ is dominated and does not belong to the weak core.

Next let $\mathcal{Z}$ be a matching described by (μ, λ) that does not belong to the weak core. Then there exist a coalition A and a matching $\mathcal{Z}'$ described by (μ', λ') that dominates $\mathcal{Z}$ on A . If $\mu'(i) = i$ for some $i \in A$, then (M3) is violated. If $\mu'(i) \neq i$ for all $i \in A$, choose some $i \in A$ and $j = \mu'(i) \neq i$. Then $\mu'(j) = i$, i and j are of opposite sex and matched in $\mathcal{Z}'$, $\lambda'(i) = \lambda'(j) =: x'$, and $(j, x') >_i (\mu(i), \lambda(i))$, $(i, x') >_j (\mu(j), \lambda(j))$, a violation of (M4). Hence (M3) or (M4) is violated which means that $\mathcal{Z}$ is unstable. ■

It immediately follows from the proposition that each stable matching is weakly Pareto optimal. However, stable matchings need not be strongly Pareto optimal. For let $M = \{m_1, m_2\}$, $W = \{w\}$, $\mathcal{X} = \{x\}$, $P(m_1) = [m_1, (w, x)]$, $P(m_2) = [m_2, (w, x)]$, $P(w) = (m_1, x), (m_2, x)$. Then $\mathcal{Z} = \{(\{m_2, w\}, x)\}$ constitutes a stable matching that is not strongly Pareto optimal.

8.3 Discussion

1. Matching with contracts. The model of this section can be reformulated in terms of matching with contracts à la Hatfield and Milgrom (2005). The set of contracts is a subset of $M \times W \times \mathcal{X}$, with generic elements $X = (m, w, x)$. However, we need not resort to this formulation. For example, due to its abstract nature, Proposition 10 can be shown directly without any further structure. But a proof along the lines of Aygün and Sönmez (2012) and Aygün and Sönmez (2013) is also possible.

2. Long-distance marriages. We can accommodate the possibility of long-distance marriages by replacing locations x by pairs of locations (x, y) where x stands for the location of the male partner and y for the location of the female partner.

3. Quasi-lexicographic preferences. Matching becomes easier and more transparent when individuals have quasi-lexicographic preferences as follows. Each person i has strict preferences for partners, given by lists $Q(i)$ with the proviso that j appears on i 's Q -list only if there is some location x so that (j, x) is acceptable to i and (i, x) is acceptable to j . For each j on i 's Q -list, there is a conditional preference list R_{ij} for locations, weakly ordering the pairs of the form (j, x) acceptable to i . Then i 's overall preference list is the array $P(i) = R_{ij_1}, R_{ij_2}, \dots, R_{ij_r}$ if $Q(i) = j_1, j_2, \dots, j_r$. For example, if $Q(m) = w_1, w_2, w_3$, $R_{mw_1} = (w_1, x_3)$, $R_{mw_2} = (w_2, x_1), (w_2, x_3)$, $R_{mw_3} = (w_3, x_1)$, then $P(m) = (w_1, x_3), (w_2, x_1), (w_2, x_3), (w_3, x_1)$. Given such quasi-lexicographic preferences, all stable matchings can be obtained by the following procedure. First, find a stable matching μ for the marriage market $(M, W; \mathbf{Q})$. For all $m \in M$ and $w \in W$ with $\mu(m) = w$, there exists $x \in \mathcal{X}$ such that (m, x) is acceptable to w and (w, x) is acceptable to m . Choose such an x that is weakly optimal for the couple – that is such that there is no $x' \in \mathcal{X}$ with $(m, x') >_w (m, x)$ and $(w, x') >_m (w, x)$ – and set $\lambda(m) = \lambda(w) = x$. Then the pair (μ, λ) represents a stable matching.

4. Limited mobility. In Sections 3 – 5, an exogenous or endogenous graph represents a restriction on feasible matches. In the current section, we do not resort to a bigraph $\mathcal{E}$ or $\mathcal{E}(s)$ with that quality. Still, we can impose restrictions on feasible locations and, consequently, feasible matches by way of preferences. For instance, (w, x) may be absent from m 's preference list because (w, x) is unacceptable to m , which in turn may stem from the fact that location x is an impossible choice for m . Suppose that, indeed, m is tied down to the particular location $x' \neq x$. Then this can be captured by the absence of all alternatives of the form (w, x) from his preference list.

9 Concluding Remarks

We represent constraints on matching by a bipartite graph. In Part 1, we perform comparative statics with respect to the exogenous constraints. Small changes in the constraints can have consequences for the entire population. Major changes in the constraints may have widespread, moderate or no effects depending on preferences. This demonstrates the crucial role of preferences.

In Part 2, we consider endogenous constraints. The examples on sabotage illustrate intriguing possibilities. However, we could not show (or disprove) existence of a Nash sabotage, which remains an open question. The analysis of choice of partners and location can be viewed as a first step towards the study of multi-dimensional matching. For instance, the literature on matching of interns and hospitals tries to allocate couples of interns to the same hospital. That is, matches in the marriage market are taken as constraints in the matching for interns and hospitals. But one can look at the situation from the opposite direction as well: The matching of interns may affect the marriage market. More generally, one would expect an interaction between the job market and the marriage market.

References

- Aygün, O. and Sönmez, T. (2012). Matching with Contracts: The Critical Role of Irrelevance of Rejected Contracts. Working Paper 804, Boston College, Department of Economics.
- Aygun, O. and Sonmez, T. (2013). Matching with Contracts: Comment. *American Economic Review*, 103(5): 2050–2051.
- Becker, G. S. (1973). A Theory of Marriage, Part I. *Journal of Political Economy* 81: 813–846. In R. Febrero and P. S. Schwartz (eds.), *The Essence of Becker*. Hoover Institution Press, Stanford, CA, 1995.
- Blair, C. (1984). Every Finite Distributive Lattice is a Set of Stable Matchings. *Journal of Combinatorial Theory*, 37(A): 353–356.
- Chakrabarti, S. and Haller, H. (2011). An Analysis of Advertising Wars. *The Manchester School*, 79: 100–124.
- Demange, G., Gale, D. and Sotomayor, M. (1987). A Further Note on the Stable Matching Problem. *Discrete Applied Mathematics*, 16(3): 217–222.
- Diestel, R. (2010). *Graph Theory*. Springer-Verlag Berlin Heidelberg, 4th ed..
- Friedman, W.A. (1998). John H. Paterson and the Sales Strategy of the National Cash Register Company, 1884 to 1922. *Business History Review*, 72(4): 552–584.
- Gale, D. and Shapley, L. S. (1962). College Admissions and the Stability of Marriage, *American Mathematical Monthly*, 69(1): 9–14.
- Gersbach, H. and Schmutzler, A. (2014). Does Globalization Create Superstars? A Simple Theory of Managerial Wages. *European Economic Review*, 71(C): 34–51.
- Gusfield, D. and Irving, R. (1989). *The Stable Marriage Problem: Structure and Algorithms*. The MIT Press, Cambridge, MA.
- Hall, P. (1935). On Representatives of Subsets. *Journal of the London Mathematical Society*, 10: 26–30.
- Hatfield, J. W. and Milgrom, P. R. (2005). Matching with Contracts. *The American Economic Review*, 95(4): 913–935.

- Holzman, R. and Samet, D. (2014). Matching of Like Rank and the Size of the Core in the Marriage Problem. *Games and Economic Behavior*, 88: 277–285.
- Jacobs, K. (1969). *Selecta Mathematica I*. HTB 49. Springer Berlin, Heidelberg, New York.
- Koopmans, T. and Beckmann, M. (1957). Assignment Problems and the Location of Economic Activities. *Econometrica*, 25(1): 53–76.
- Lovász, L. and Plummer, M. D. (1986). *Matching Theory*. North-Holland Amsterdam, New York.
- Lubin, G. (2011). Campbell’s Soup is Still Reeling from the 2008 MSG Attack Ads. Available at: <http://www.businessinsider.com.au/campbells-soup-wars-citi-2011-7>
- Maak, W. (1936). Eine neue Definition der fastperiodischen Funktionen. *Abhandlungen aus dem Mathematischen Seminar Hamburg*, 11: 240–244.
- Roth, A. and Sotomayor, M. (1990). *Two-Sided Matching Theory: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, 1990.
- Shapley, L. S. and Shubik, M. (1972). The Assignment Game I: The Core. *International Journal of Game Theory*, 1(1), 111–130.
- von Weizsäcker, C. C. (1971). Notes on Endogenous Change of Tastes. *Journal of Economic Theory*, 3(4), 345–372.
- Weyl, H. (1949). Almost Periodic Invariant Vector Sets in a Metric Vector Space. *American Journal of Mathematics*, 71(1), 178–205.