Targeted Search in Matching Markets[*]

Anton Cheremukhin,   Paulina Restrepo-Echavarria,
Antonella Tutino

March 2, 2017

Abstract

We propose a search and matching model where agents can choose who to meet but have to pay a search cost to locate potential matches more accurately. The model features a tension between an agent's desire to find a more productive match and to maximize the odds that interest is mutual. This tension drives a wedge between the shape of sorting patterns and the shape of the underlying match payoff function allowing for better empirical identification of preferences. We show the empirical relevance of our model by applying it to the U.S. marriage market.

JEL: E24, J64, C78, D83.

Keywords: Matching, sorting, assignment, search.

1 Introduction

The search for a partner—whether in business or in personal life—includes both productive and strategic considerations: People seek productive partnerships that maximize their pay-offs, but also search strategically by considering the likelihood a potential partner will also be interested in them. As a result, we observe that matching is not perfectly assortative nor purely random. For instance, the empirical marriage literature documents an abundance of matches between inferior and superior partners.[1] Understanding why this is the case is important for understanding the roles played by preferences and frictions in the matching process, and for evaluating losses from mismatch.

One reason for apparent mismatch may be that the econometrician does not observe all the match-relevant characteristics.[2] However, even studies with exhaustive information about the qualities of potential matches and their choice patterns leave a large part of the variation in the data unexplained.[3]

Another important reason is search frictions. Matches between inferior and superior types may form if one of the partners cannot afford to "wait for the best" and decides to "settle for the rest".[4] Who marries whom is also influenced by who meets whom.[5] Although circumstances are a major driver of the meeting process, whom to meet is also a choice. To our knowledge, the existing literature is silent on how to model this choice. Instead, it is common to either postulate an exogenous random meeting process or assume

---

[1]See survey by Chiaporri and Salanie (2015) and references therein.
[2]See Choo and Siow (2006) and Galichon and Salanie (2012).
[3]See Hitsch et. al. (2010).
[4]See Eeckhout (1999), Shimer and Smith (2000), Adachi (2003).
[5]See e.g. Belot and Francesconi (2013).

that meeting the best partner is costless.

In this paper we propose a parsimonious way to model the choice of whom to meet, that captures the salient features of the meeting process. One reason that people cannot start their search by simply contacting the best partners is that they are uncertain who those are. However, people can choose to put some effort into finding out about the options and locating the best partners. Empirical studies of discrete choice suggest that people respond to incentives by putting more effort and locating best options with a larger probability when stakes are higher.[6] The tendency to choose options with a higher payoff more often has also been extensively documented for marriage market choices.[7] To accommodate this empirical pattern, we build our modeling strategy on the growing stochastic discrete choice literature. We assume that agents choose who to contact in a probabilistic way. The strategies chosen by agents are their distributions of interest. Agents are able to target one or several potential partners by contacting them with higher probability. However, increasing the likelihood of contacting the best partner involves cognitive effort. We capture this intuition by associating a proportional cost with a measure of distance between an "uninformed" strategy and the agent's strategy of choice. We borrow the specification of cost from the burgeoning literature on discrete choice under information frictions.[8] With this approach, our model naturally captures the intuition that agents choose to contact more often the potential partners that yield a higher expected payoff.

Another reason for contacting a potential partner is confidence that interest is likely to be mutual, while partners that are unlikely to reciprocate

---

[6]See the large literature starting with Hey and Orme (1994).
[7]See Fisman et. al. (2006) and Hitsch et. al. (2010).
[8]See Cheremukhin et. al. (2015) and Matejka and McKay (2015).

are worth excluding. Thus, people act strategically not only when deciding whether to form a match or wait for a better option, but also when choosing who to contact. Our model captures this second strategic motive, as agents' equilibrium strategies respond positively both to the utility they derive from a match, and the likelihood that interest will be mutual. Thus, our model endogenizes the randomness of the meeting process by formalizing the fact that when looking for a match agents have two incentives: first, to maximize their actual payoffs (which we refer to as the productive motive), and second, to maximize the odds of forming a match (which we refer to as the strategic motive).

We first present a bare-bones one-shot model that has the minimal ingredients to showcase the mechanism. We define a matching equilibrium of the model as the pure-strategy Nash equilibrium between agents' probabilistic strategies. This solution concept is less restrictive than the stability requirement coming from cooperative games. Nonetheless, we prove existence of equilibirum, derive the necessary conditions characterizing it, and show sufficient conditions for its uniqueness. Furthermore, we show that the equilibria that emerge from a positive and finite cost are inefficient relative to the constrained Pareto allocation, although the outcomes are constrained efficient in both limiting cases.

Our theoretical results are invaluable for building intuition. The meeting rates in the model are determined by the interaction between the productive and strategic motives. The relative strength of the two motives depends on the search cost. When the cost is low, locating the best partner requires little effort, and the likelihood of reciprocation increases. In this case, the reciprocity of interest is the paramount determinant of who meets whom, while the importance of joint productivity may be undermined by the lack

3

of mutual interest. This case supports multiple equilibrium arrangements of mutual interest, but only a subset of these arrangements approaches the set of stable allocations in the frictionless limit.

When it is costly to locate the best partner, the likelihood of contacting an inferior partner increases. Payoffs become the driving force behind who meets whom. In the unique equilibrium, every agent's strategy is to chase the partner that would yield the highest payoff, while the impact of mutuality of interest is minimal.

This property can be used for empirical identification of agents' preferences. In particular, the literature describes "horizontal" (attraction to similar agents) and "vertical" (attraction based on a commonly agreed-upon ranking) preferences,[9] but finds it hard if not impossible to distinguish between these cases empirically.[10] Indeed, both cases lead to identical assortative stable matching in the frictionless case. In contrast, the equilibrium matching rates predicted by our model differ markedly for these two cases. When preferences are "horizontal", the strategic and productive motive pull agents in the same direction, as look-alikes both get the highest payoff from each other and their interest is also more likely to be mutual. In this case, a stochastic version of assortative matching is preserved in equilibrium. However, in the case of vertical preferences, there is common agreement on the ranking of agents and everybody tries to chase the top type despite the lower likelihood of mutual interest. The productive and the strategic motive pull in opposite directions. This case gives rise to a novel and surprising equilibrium pattern that reminisces neither positive nor negative assortative matching. We call it the mixing equilibrium.[11]

---

[9]See Banerjee et. al. (2013) and references therein.

[10]See Hitsch et. al. (2010) and related studies.

[11]Our taxonomy of equilibria in this case follows that of Burdett and Coles (1999).

This novel matching pattern can be used to distinguish empirically between horizontal and vertical preferences. We analyze how preferences can be identified from the data on matching rates and from the data on contact rates if those are observed. To show the empirical relevance of these theoretical results, we take our model to the data. We find that the model does a very good job rationalizing the observed matching rates in the U.S. marriage market based on income, age, and education, and for these three cases we estimate the underlying payoffs through the lens of our model. Our empirical results suggest that strategic interactions can be used to recover preferences, which may indeed differ markedly from observed matching patterns. Ignoring this difference may result in misleading estimates of preferences and, hence, of losses from mismatch.

Our paper effectively blends two sources of randomness used in the empirical literature. The first source is a search friction with uniformly random meetings and impatience, as in Shimer and Smith (2000). The second approach introduces unobserved characteristics as a tractable way of accounting for the deviations of the data from the stark predictions of the frictionless model, as in Choo and Siow (2006) and Galichon and Salanie (2012). We introduce a search friction into the meeting process by endogenizing agents' choice of whom to contact. We build on the discrete choice rational inattention literature—i.e., Cheremukhin, Popova, and Tutino (2015) and Matejka and McKay (2015)—that derives multinomial logit decision rules as a consequence of cognitive constraints that capture limits to processing information. Therefore, the equilibrium matching rates in our model have a multinomial logit form similar to that in Galichon and Salanie (2012). Unlike Galichon and Salanie, the equilibrium of our model features strong interactions between agents' contact rates driven entirely by their conscious choices, rather

than by some unobserved characteristics with fixed distributions.

The search and matching literature has seen multiple attempts to produce intermediate degrees of precision with which agents meet their best choices. In particular, Menzio (2007) and Lester (2011) nest directed search and random matching to generate outcomes with an intermediate degree of randomness.[12] Our paper produces equilibrium outcomes in between uniform random matching and the frictionless assignment, without nesting these two frameworks. Although the directed search literature, such as Eeckhout and Kircher (2010) and Shimer (2005),[13] technically involves a choice of who to meet, the choice is degenerate - directed by signals from the other side. The key friction in directed search is the congestion externality where agents on one side of the market compete with each other for matching with specific agents on the other side. Congestion slows down matching and can produce mismatch but does not distort sorting patterns compared with frictionless assignment. The search friction in our model results in mis-coordination between agents on opposite sides of the market, which leads to detectable distortions of sorting patterns.

The paper proceeds as follows: Section 2 describes the model and derives the theoretical results. In Section 3 we provide an extensive discussion of properties of equilibria and how they can be used to identify preferences. We apply the model to the U.S. marriage market data in Section 4. Section 5 states some final remarks.

---

[12]Also, see Yang's (2013) model of "targeted" search that assumes random search within perfectly distinguishable market segments.

[13]See Chade, Eeckhout and Smith (2016) for a neat summary of this literature.

2 Model

We build on the frictionless matching environment of Becker (1973), where males and females are heterogeneous in their type and all are searching for a match. Both males and females know the distribution and their preferences over types on the other side of the market, but there is noise—agents cannot locate potential partners with certainty. However, they can pay a search cost to help locate them more accurately.

We model this by assuming that each agent chooses a probability distribution over types. This distribution reflects the likelihood of contacting a particular agent on the other side. A more targeted search, or a probability distribution that is more concentrated on a particular group of agents is associated with a higher cost as the agent needs to exert more effort in deciding who to contact. The probability distribution needs to satisfy two properties: 1) By the nature of the choice between a finite number of options, the distribution must be discrete; and 2) for strategic motives to play a role, agents should be able to vary each element of the distribution and consider small deviations of each element in response to changes in the properties of the options. Hence, this probability distribution cannot be confined to a specific family of distributions.

The choice of functions in economics that satisfy these requirements is very limited. We use the Kullback-Leibler divergence (relative entropy) as the measure of search effort. This specification accommodates both full choice of a distribution and a discrete choice problem. In addition, it turns out that, in our specific case of a choice among discrete options, this specification enhances tractability and leads to closed-form solutions. Specifically, the solution has the form of a multinomial logit that is well understood and already widely used in empirical studies of discrete choice environments and

7

of the marriage market.

After choosing their optimal probability distribution over types, both males and females simultaneously make a single draw from their distributions. If the draw is reciprocated, a match is formed if it is mutually beneficial and the output from the match is split between the two parties.

2.1 Environment

There are F females indexed by $x \in \{1, ..., F\}$ and M males indexed by $y \in \{1, ..., M\}$. Both males and females are heterogeneous in types and are actively searching for a match.[14] A match between female x and male y generates a payoff Φ_{xy}.[15] If a male and a female match, the payoff is split between them. We normalize the outside option of both to zero. We denote the payoff appropriated by the female ε_{xy} and the payoff appropriated by the male η_{xy}, such that $\Phi_{xy} = \varepsilon_{xy} + \eta_{xy}$.[16] The payoff and the split generated by any potential (x, y) match are known ex-ante to female x and male y.

Each female chooses a discrete probability distribution, $p_x(y)$, which reflects the probability with which female x will target male y (seek him out). Each female x rationally chooses her strategy while facing a trade-off between a higher payoff and a higher cost of searching. Likewise we denote the strategy of a each male $q_y(x)$. It represents the probability of a male y targeting a female x. Each agent can vary and choose each element of their distribution. Placing a higher probability on any particular potential match, implies that the agent choosing the distribution has exerted more search effort, will target a potential partner more accurately and hence, will have a

---

[14]The extension to multiple identical agents of each type is straightforward and is not discussed in the paper.

[15]Note that we do not place any restrictions on the payoff function.

[16]For simplicity we assume that payoffs are fixed. We discuss the role played by (non)transferrability in Section 3.1.

higher probability of matching with them.

A female's total cost of searching is given by $c_x\left(\kappa_x(p_x(y))\right)$. This cost is a function of the search effort, κ_x, and hence of the probability distribution, $p_x(y)$, chosen by female x. Likewise, we denote a male's cost of searching by $c_y\left(\kappa_y(q_y(x))\right)$, where the cost is a function of the search effort, κ_y, and hence of the probability distribution, $q_y(x)$, chosen by male y.

Figure 1 illustrates the strategies of males and females. Consider a female $x = 1$. The solid arrows show how she assigns a probability $p_1(y)$ of targeting each male y. Similarly, dashed arrows show the probability $q_1(x)$ that a male $y = 1$ assigns to targeting a female x. Once these are selected, each male and female will make one draw from their respective distribution to determine which individual they will actually contact. A match is formed between male y and female x if and only if: 1) according to the female's draw of y from $p_x(y)$, female x contacts male y; 2) according to the male's draw of x from $q_y(x)$, male y also contacts female x; and 3) their payoffs are non-negative.
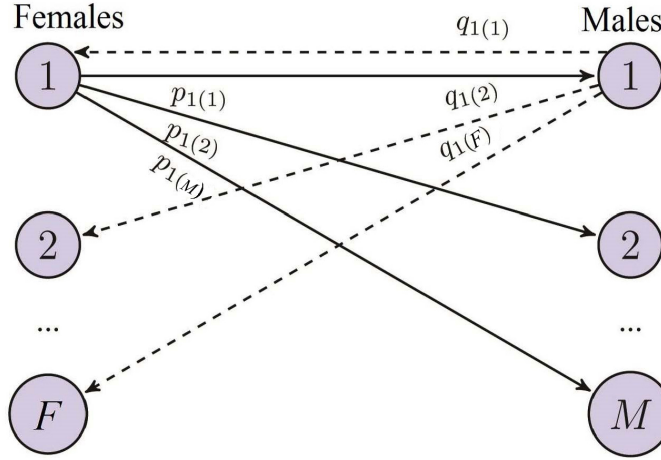

Figure 1. Strategies of males and females

Since negative payoffs lead to de facto zero payoffs due to the absence of

a match, we can assume that all payoffs are non-negative:

$$\Phi_{xy} \geq 0, \qquad \varepsilon_{xy} \geq 0, \qquad \eta_{xy} \geq 0.$$

Each female x chooses a strategy $p_x(y)$ to maximize her expected net payoff:

$$Y_x = \max_{p_x(y)} \sum_{y=1}^{M} \varepsilon_{xy} q_y(x) p_x(y) - c_x(\kappa_x(p_x(y))).$$

The female gets her expected return from a match with male y net of the cost of searching. The probability of a match between female x and male y is given by the product of the distributions $q_y(x) p_x(y)$. Note that in equilibrium the matching rate that female x faces from male y equals male y's strategy $q_y(x)$. As matching rates are equilibrium objects, they are assumed to be common knowledge to participating parties.

The cost function is given by $c_x(\kappa_x) = \theta_x \kappa_x$, where θ_x is the marginal cost of search. Here, we are using the linear cost function for simplicity, but all of our proofs will hold for more general cost functions. As mentioned earlier, κ_x reflects search effort and needs to accommodate the full choice of a discrete distribution. One function that satisfies these requirements is the following:[17]

$$\kappa_x = \sum_{y=1}^{M} p_x(y) \ln \frac{p_x(y)}{1/M}, \tag{1}$$

where $p_x(y)$ must satisfy $\sum_{y=1}^{M} p_x(y) = 1$ and $p_x(y) \geq 0$ for all y.

---

[17]In the model of information frictions used in the rational inattention literature, κ_x, would represent the relative entropy between a uniform prior $\{1/M\}$ over males and the posterior strategy, $p_x(y)$. This definition is a special case of Shannon's channel capacity where information structure is the only choice variable (See Thomas and Cover (1991), Chapter 2). See also Cheremukhin et al (2015) for an application to stochastic discrete choice with information costs.

Note that, κ_x, is increasing in the distance between a uniform distribution $\{1/M\}$ over males and the chosen strategy, $p_x(y)$. If an agent does not want to exert any search effort, she can choose a uniform distribution $p_x(y) = \frac{1}{M}$ over types, the effort involved in search is zero, and her search is random. As she chooses a more targeted strategy the distance between the uniform distribution $\{1/M\}$ and her strategy $p_x(y)$ is greater, increasing κ_x and the overall cost of searching, and her search will be less random. By increasing search effort agents bring down uncertainty about the location of a prospective match, which allows them to target their better matches more accurately.

Similarly, male y chooses his strategy $q_y(x)$ to maximize his expected payoff:

$$Y_y = \max_{q_y(x)} \sum_{x=1}^{F} \eta_{xy} p_x(y) q_y(x) - c_y(\kappa_y(q_y(x))),$$

where

$$\kappa_y = \sum_{x=1}^{F} q_y(x) \ln \frac{q_y(x)}{1/F}, \tag{2}$$

and $q_y(x)$ must satisfy $\sum_{x=1}^{F} q_y(x) = 1$ and $q_y(x) \geq 0$ for all x.

2.2 Matching equilibrium

Definition 1. *A* matching equilibrium *is a set of strategies of females,* $\{p_x(y)\}_{x=1}^{F}$, *and males,* $\{q_y(x)\}_{y=1}^{M}$, *that simultaneously solve problems of males and females.*

The equilibrium of the matching model can be interpreted as a pure-strategy Nash equilibrium of a strategic form game. In what follows we shall apply the results for concave n-person games from Rosen (1965). The game consists of the set of players, the set of actions and the player's payoffs. The

set of players is given by $\mathcal{I} = \{x \in \{1, ..., F\}, y \in \{1, ..., M\}\}$. The set of actions $s \in S$ is given by the cartesian product of the sets of strategies of females $p_x(y) \in S_x$ and males $q_y(x) \in S_y$, where

$$S_x = \left\{ p_x(y) \in R^M, p_x(y) \geq 0, \sum_{y=1}^{M} p_x(y) \leq 1 \right\},$$

$$S_y = \left\{ q_y(x) \in R^F, q_y(x) \geq 0, \sum_{x=1}^{F} q_x(x) \leq 1 \right\}.$$

The payoffs $u_i(s) = \{Y_x(s), Y_y(s)\}$ are defined as follows:

$$Y_x(p_x(y), q_y(x)) = \sum_{y=1}^{M} \varepsilon_{xy} q_y(x) p_x(y) - c_x \left(\sum_{y=1}^{M} p_x(y) \ln \frac{p_x(y)}{1/M} \right),$$

$$Y_y(q_y(x), p_x(y)) = \sum_{x=1}^{F} \eta_{xy} p_x(y) q_y(x) - c_y \left(\sum_{x=1}^{F} q_y(x) \ln \frac{q_y(x)}{1/F} \right).$$

Theorem 1. *A matching equilibrium exists.*

Proof. Note that the strategy set of each player is a unit simplex, and therefore a non-empty, convex and compact set. For a pure-strategy Nash equilibrium to exist, each payoff function $u_i(s)$ needs to be continuous in the strategies s, and $u(s_i, s_{-i})$ needs to be quasiconcave in s_i. Indeed, under the assumption that cost functions are continuous, non-decreasing and (weakly) convex, the payoff functions are continuous and concave in the own strategies of players. For the case of a linear cost function these restrictions are trivially satisfied.

$\square$

To show uniqueness, we need to introduce some additional notation. Note that for each player $i \in \mathcal{I}$ the strategy set can be represented as $S_i = \{s_i \in R^{m_i}, h_i(s_i) \geq 0\}$ where h_i is a concave function. Indeed in our case, the

12

functions $h_x(p_x(y)) = \left[p_x(1), ..., p_x(M), 1 - \sum_{y=1}^{M} p_x(y)\right]$ and $h_y(q_y(x)) = \left[q_y(1), ..., q_y(F), 1 - \sum_{x=1}^{F} q_y(x)\right]$ are concave. Following Rosen, define the gradient $\nabla u(s) = [\nabla_1 u_1(s), ..., \nabla_m u_m(s)]^T$ and Hessian: $U(s) = \nabla_i \nabla_j u_i(s)$. Then, if the constraints h_i are concave and the symmetrized Hessian $U(s) + U^T(s)$ is negative definite for all $s \in S$ then the payoff functions are diagonally strictly concave for $s \in S$. We can then use the result that if h_i are concave functions, if there exist interior points $\tilde{s}_i \in S_i$ such that $h_i(\tilde{s}_i) > 0$, and if the payoff functions are diagonally strictly concave for all $s \in S$ then the game has a unique pure strategy Nash equilibrium.

Theorem 2. *The matching equilibrium is unique if*

a) cost functions are non-decreasing and convex;

b) $\left.\frac{\partial c_x(\kappa_x)}{\partial \kappa_x}\right|_{p_x^*(y)} = \theta_x > \varepsilon_{xy} p_x^*(y);$

c) $\left.\frac{\partial c_y(\kappa_y)}{\partial \kappa_y}\right|_{q_y^*(x)} = \theta_y > \eta_{xy} q_y^*(x).$

Proof. If the cost functions $c(\kappa)$ are (weakly) increasing and (weakly) convex in κ, then the payoffs of all males and females are continuous and also concave in their strategies. Assuming that the cost functions are twice continuously differentiable functions, the Hessian of this game is the matrix of all second derivatives. The diagonal elements are all non-positive, consistent with concavity of the payoffs:

$$\frac{\partial^2 Y_x}{\partial p_x \partial p_x} = -\frac{\partial c_x}{\partial \kappa_x}(\kappa_x)\frac{1}{p_x(y)} - \frac{\partial^2 c_x}{\partial \kappa_x \partial \kappa_x}(\kappa_x)\left(1 + \ln\frac{p_x(y)}{1/M}\right)^2 \le 0,$$

$$\frac{\partial^2 Y_y}{\partial q_y \partial q_y} = -\frac{\partial c_y}{\partial \kappa_y}(\kappa_y)\frac{1}{q_y(x)} - \frac{\partial^2 c_y}{\partial \kappa_y \partial \kappa_y}(\kappa_y)\left(1 + \ln\frac{q_y(x)}{1/F}\right)^2 \le 0.$$

The off-diagonal elements are all non-negative:

$$\frac{\partial^2 Y_x}{\partial p_x \partial q_y} = \varepsilon_{xy} \ge 0,$$

13

$$\frac{\partial^2 Y_y}{\partial q_y \partial p_x} = \eta_{xy} \geq 0.$$

The remaining cross-derivatives are all zero. Note also that the Hessian is itself symmetric so there is no need to symmetrize it. To guarantee that the Hessian is negative definite, we require the following "Diagonal dominance" conditions:

$$\left| \frac{\partial^2 Y_x}{\partial p_x \partial p_x} \right| > \left| \frac{\partial^2 Y_x}{\partial p_x \partial q_y} \right|,$$
$$\left| \frac{\partial^2 Y_y}{\partial q_y \partial q_y} \right| > \left| \frac{\partial^2 Y_y}{\partial q_y \partial p_x} \right|.$$

Diagonal dominance conditions postulate that diagonal elements of the Hessian are larger in absolute value than any off-diagonal elements, which in turn guarantees that the Hessian of the game is negative definite. It is clear that when the cost functions are linear, these conditions simplify to $\theta_x \frac{1}{p_x(y)} > \varepsilon_{xy}$ and $\theta_y \frac{1}{q_y(x)} > \eta_{xy}$. While Rosen's version requires that these conditions hold globally for all $s \in S$ which would imply $\theta_x > \varepsilon_{xy}$ and $\theta_y > \eta_{xy}$, these conditions could be relaxed to require diagonal dominance to be satisfied only along the equilibrium path. For this we note that since the constraints are given by unit simplexes (for which index equals one, and every KKT point is complementary and non-degenerate) we can invoke the generalized Poincare-Hopf index theorem of Simsek, Ozdaglar, and Acemoglu (2007) which in this case implies that the equilibrium is unique if the Hessian is negative-definite at the equilibrium point. Thus, the equilibrium is unique if diagonal dominance conditions hold only along the equilibrium path, i.e. if conditions (b) and (c) are satisfied. □

Note that the assumptions we make to prove uniqueness are by no means restrictive. The assumption that cost functions are non-decreasing and convex is a natural one. The additional "diagonal dominance" conditions in our

case can be interpreted as implying that the search cost should be sufficiently high for the equilibrium to be unique. If these conditions do not hold, then there can be multiple equilibria. This is a well-known outcome of the assignment model, which is a special case of our model under zero search costs. In a frictionless environment, the multiplicity of equilibria is eliminated by requiring that the matching be "stable", a solution concept from cooperative games requiring that there is no profitable pairwise deviation. In our framework, checking for pairwise deviations would require that all males know the location of all females and vice versa. Since locating agents is costly in our model, we use the Nash equilibrium solution concept, which implies that the equilibrium outcome generically does not satisfy "stability."

Under the assumptions on the cost functions made earlier we can also obtain a characterization result. The derivatives of the constrained payoff functions with respect to own strategies are:

$$\frac{\partial Y_x}{\partial p_x} = \varepsilon_{xy} q_y\left(x\right) - \frac{\partial c_x}{\partial \kappa_x}\left(\kappa_x\right)\left(1 + \ln \frac{p_x\left(y\right)}{1/M}\right) - \lambda_x,$$

$$\frac{\partial Y_y}{\partial q_y} = \eta_{xy} p_x\left(y\right) - \frac{\partial c_y}{\partial \kappa_y}\left(\kappa_y\right)\left(1 + \ln \frac{q_y\left(x\right)}{1/F}\right) - \lambda_y.$$

When cost functions are non-decreasing and convex, it is easy to verify that first-order conditions are necessary and sufficient conditions for equilibrium. Rearranging the first-order conditions for males and females, we obtain

$$p_x^*\left(y\right) = \exp\left(\frac{\varepsilon_{xy} q_y^*\left(x\right)}{\theta_x}\right) \bigg/ \sum_{y'=1}^{M} \exp\left(\frac{\varepsilon_{xy'} q_{y'}^*\left(x\right)}{\theta_x}\right),$$

$$q_y^*\left(x\right) = \exp\left(\frac{\eta_{xy} p_x^*\left(y\right)}{\theta_y}\right) \bigg/ \sum_{x'=1}^{F} \exp\left(\frac{\eta_{x'y} p_{x'}^*\left(y\right)}{\theta_y}\right). \tag{3}$$

These necessary and sufficient conditions for equilibrium cast the optimal strategy of female x and male y in the form of a best response to optimal strategies of males and females, respectively.

3   Discussion

3.1   Properties of matching equilibria

The result of Theorem 2 is intuitive. Recall that there are two motives for female x to target male y: the productive and the strategic motive. The payoff of a female depends on the product of the portion she appropriates from the output of the match and the probability of reciprocation. While her private payoff does not depend on equilibrium strategies, the strategic motive does.

When the search cost, θ, is very low, females (and males) are able to place a high probability of targeting one counterparty and exclude all others. It does not matter what portion of the payoff female x will get from a match with male y if the male places a low probability on female x. In the extreme, any pairing of agents is an equilibrium since no one has an incentive to deviate from any mutual reciprocation. The strategic motive dominates and multiplicity of equilibria is a natural outcome. As the search costs go to zero, targeting strategies become more and more precise. In the limit, in every equilibrium each female places a unit probability on a particular male, and that male responds with a unit probability of considering that female. Each equilibrium of this kind implements a matching of the classical assignment problem (not all of them are stable).

As θ increase, probability distributions become less precise, as it is increasingly costly to target a particular counterparty. That is, the search costs dampen the strategic motive and the productive motive plays a bigger role. At some threshold level of θ the strategic motive is dampened enough that all agents will choose probabilities primarily seeking a match with a higher payoff. This level of costs is characterized by the "diagonal dominance" conditions of Theorem 2. Agents require the strategic motive,

characterized by the off-diagonal element of the Hessian of the game, to be lower than the productive motive, captured by the diagonal element. Above the threshold, the unique equilibrium has the property that each agent places a higher probability on the counterparty that promises a higher payoff, i.e., the productive motive dominates. When search costs go to infinity, optimal strategies of males and females approach a uniform distribution. This unique equilibrium implements the standard uniform random matching assumption extensively used in the literature. Thus, the assignment model and the random matching model are special cases of our targeted search model, when θ is either very low or very high.

Equilibrium conditions (3) also have an intuitive interpretation. They predict that the higher the female's private gain from matching with a male, the higher the probability of targeting that male. Males are naturally sorted in each female's strategy by the probability of the female targeting each male. The strategies of males have the same properties due to the symmetry of the problem. Theorem 2 predicts that an increase in θ reduces the interaction between search strategies of females and males. Once θ is sufficiently high, the intersection of best responses leads to a unique equilibrium. Note that, by the nature of the index theorem used in the proof of uniqueness, it is enough to check diagonal dominance conditions locally in the neighborhood of the equilibrium. There is no requirement for them to hold globally. This suggests a simple way of computing the unique equilibrium. We first need to find one solution to the first-order conditions (3) and then check that diagonal dominance conditions are satisfied.

In appendix A we analyze whether an equilibrium of the model is efficient from the point of view of a constrained planner. We find that except for the two extreme cases, the equilibria of the model are socially inefficient. While

it is socially optimal for both females and males to consider the total payoff of the match, in the decentralized equilibrium they consider only their private payoffs. This result is reminiscent of goods with positive externalities where the producer undersupplies the good if she is not fully compensated by the marginal social benefits that an additional unit of the good would provide to society. In our model, additional search effort exerted by an individual male or female has a positive externality on the whole matching market. In appendix B we show that the constrained optimum is possible to achieve if only one side of the market searches actively.

In appendix C we consider an extension of the two-sided matching model to a repeated setting. Following Eeckhout (1999) and Adachi (2003) we assume that at the moment that male y and female x meet, each of them has an additional decision to make. Each agent may choose to form a match and receive the corresponding share of the surplus, or refuse to form a match and wait for a better potential partner in future periods if their continuation value is higher than the payoff from matching with the proposed partner. The continuation value is assumed to be simply the expected payoff from matching in the future discounted at the rate ρ, a patience parameter.

When the agents are unable to distinguish partners until they meet, i.e. when the cost parameter θ approaches infinity, we obtain the Adachi (2003) model. In that case, if the patience parameter ρ approaches 1, the model replicates the frictionless matching outcome as agents are able to wait as long as necessary to meet their best match. Similarly, when agents cannot wait and match everybody that they meet, i.e. the patience parameter ρ is set to zero, we obtain our baseline one-shot model. In that case, if agents are nonetheless able to perfectly distingish among potential partners, i.e. the parameter θ approaches zero, the model, with a refinement permitting only

stable matchings, also reproduces the frictionless matching outcome.

The repeated game with patience is instructive as it highlights two independent sources of search frictions: the costs of waiting and the costs of distinguishing among agents. According to Smith et al. (1999), search costs are divided into external and internal costs. External costs include the monetary costs of searching and contacting partners as well as the opportunity costs of the time spent on searching. These costs are captured by the parameter ρ in the repeated model. Internal costs include the mental effort associated with the search process, sorting the incoming information, and integrating it with what the agent already knows. Modeling the internal costs is the novel feature of our model. Internal costs are captured by parameter θ which describes the agents' ability to evaluate available information, depending on intelligence, prior knowledge, education and training. The properties of the extended model highlight that both internal and external costs of search are necessary to obtain outcomes where superior agents are matched with inferior agents in equilibrium: the agents need to be both reasonably impatient and unable to perfectly distinguish among potential partners. Although the two types of frictions are quite different in nature, we find that they reinforce each other: if agents can distinguish their best matches better, the equilibrium likelihood of meeting is higher, which increases the continuation value of waiting, just like an increase in patience.

This extension also highlights two distinguishing features of our model. First, it emphasizes the difference between the choice of whom to meet constrained by cognitive costs, and the choice of whether to form a match or keep looking for a better one constrained by the physical costs. Ours is an explicit model of how agents choose who to meet. Second, the extended model makes clear the source of the difference between the TU (transferable util-

ity) and the NTU (non-transferable utility) cases. If agents are able to reject potential partners that are not good enough, then it is important whether those potential partners can offer a larger share of the surplus in return for forming a match. The more impatient agents are, the smaller the difference between the TU and NTU cases. In our one-shot model the TU case and the NTU case are identical as the continuation values are zero and all matches are viable.

3.2   Implications for Sorting

To better understand the effect of the productive and strategic motives on equilibrium strategies and matching rates it is useful to consider simple examples of payoffs. Let us consider a matching market where there are just two males and two females, with types labeled, high (H) and low (L). Let us also consider two specific cases of the form of the payoff function, which in the literature are often referred to as horizontal and vertical preferences.[18]

Case one: The high type female is better off with a high type male, and the low type female is better off with a low type male. The same property is true for males. We shall generally refer to a payoff function where for each type the best option on the other side is different - as the case of horizontal preferences. Case two: Both females prefer the high type male, and both males prefer the high type female. We shall generally refer to a payoff function for which everyone's best option is the same type as the case of vertical preferences. These definitions only place restrictions on the structure of agetns' best options, and are therefore less restrictive than existing definitions in the literature.

In the case of horizontal preferences, the strategic and the productive

---

[18]See e.g. Hitsch et al (2010) and Herrenbrueck et al (2016).

motives are aligned. The productive motive points all agents in different directions - towards their best options, and the strategic motive ensures that the same agent that implies a higher payoff is also the one that is more likely to reciprocate (because agents have no incentive to compete for the same match). However, in the case of vertical preferences, the productive motive points all agents in the same direction, while the strategic motive tends to coordinate agents on paying attention to those whom their competitors are less likely to consider, to maximize the odds of finding a match. Thus, there is a conflict between the two motives as they pull intentions in different directions.

If preferences are horizontal, and the search costs are low, our model can have two different equilibrium patterns. The first pattern is where the high type is more likely to target the high type and the low type to target the low type (HH, LL). This is the case of positive assortative matching (PAM). The second pattern is when the high type is more likely to target a low type, because the low type is more likely to reciprocate (HL, LH). This is the case of negative assortative matching (NAM). However, if search costs are high, only the PAM equilibrium survives because the productive motive dominates.

If preferences are vertical, and search costs are low, in addition to the PAM and NAM equilibria, there is a third equilibrium pattern, which we call a *mixing* equilibrium. In the mixing equilibrium, both females target the high type male, and both males target the high type female. Moreover, for high enough search costs, the unique equilibrium has the mixing pattern, while the PAM and NAM equilibria disappear. These patterns are illustrated in Figure 1.

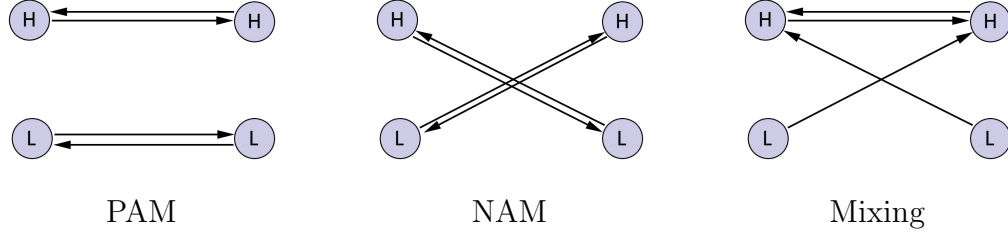This last result is in stark contrast with the literature on optimal assign-

Figure 1: Three types of equilibria and their sorting patterns
Note: We show by an arrow the direction in which each agent places the highest probability.

ment, which predicts a PAM equilibrium as the only stable outcome for either horizontal or vertical preferences. The prediction of the assignment model is driven by the strategic motive. If search costs are low, the high types are only interested in each other, so it makes no sense for the low types to target the high types as, despite a higher potential payoff, the chance their interest will be reciprocated is zero. However, when search costs are high enough, the strategic motive is dampened to the extent that the productive motive starts to play a dominant role. The productive motive instructs people to place a higher probability on the type that promises a higher payoff. Hence, the unique mixing equilibrium.

This basic intuition has important implications for empirical inference. If the productive and strategic motives are perfectly aligned, as they are for horizontal preferences, then the shape of the equilibrium matching pattern is very similar to the shape of the payoff function. The presence of a conflict between these motives, as in the case of vertical preferences, drives a wedge between the shape of the payoff and the shape of the matching rates. The conflict between motives creates a large number of competing agents that would be able to compensate for the lower payoff by a higher probability of reciprocation. Therefore, the pattern of who meets whom will differ substantially from the pattern of who would be better off with whom.
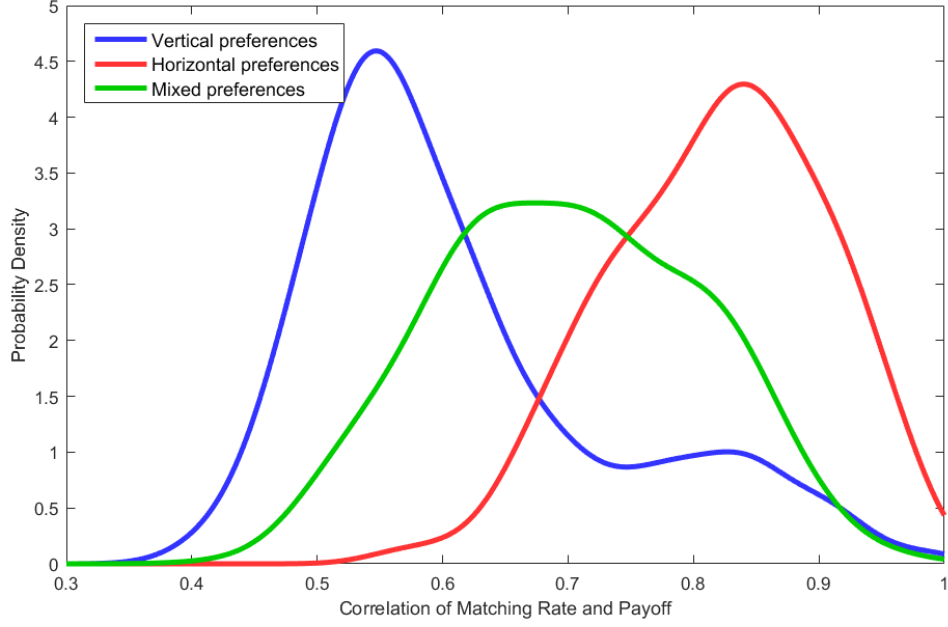
Figure 2: Correlation of Matching rate and Payoff

To quantify this difference, we run a set of Monte Carlo simulations and compute the correlation between the equilibrium matching rate and the underlying payoff function. For the Monte Carlo simulations, we assume three males and three females and draw each element of the 3-by-3 payoff matrix from a uniform distribution. We make 25,000 draws. We then find all equilibria and corresponding matching rates for each draw of the payoff function. For all draws, we compute the correlation between the matrix of equilibrium matching rates and the payoff matrix. In Figure 2 we show the probability density functions of correlations for three classes of payoff functions: vertical preferences, horizontal preferences, or no clear preference pattern.

We find that, indeed, in the case of vertical preferences the correlation is significantly lower than that in the case of horizontal preferences. The intermediate shapes of payoffs generate intermediate values of the correlation.

Thus, when our model is the true data-generating process, the conflict between the productive and strategic motive drives a substantial wedge between the shape of the underlying payoff function and the shape of the matching rate. Consequently, the empirical researcher could easily arrive at wrong conclusions about the shape of the underlying payoff and the optimal frictionless allocation by simply looking at the shape of the matching rates. As we shall discuss at the end of the empirical section, this is indeed what workhorse models of the marriage market do.

To put this result in context, we note that both random matching models a la Shimer and Smith (2000), and directed search models a la Eeckhout and Kircher (2010) can produce a substantial wedge between the shape of the payoff function and the shape of the matching rates. In the case of random matching the distribution is uniform, while in the case of directed search matching is fully assortative. Matching patterns in both of these cases are accommodated by our model under extreme (very high or very low) values of search costs. Our model also spans the continuum of matching patterns in between these two extremes.

To show that the wedge between the matching pattern and the payoff function is indeed present in the data and empirically relevant, in the empirical section, we explore three prominent examples of matching patterns in the marriage market. We show that, when viewed through the lens of our model, they exhibit strong vertical preferences. Also, we observe a substantial wedge between the shape of the underlying payoff function and the matching rate.

3.3 Identification

Identifying preferences empirically, in particular, distinguishing between horizontal and vertical preferences, is hard because both cases lead to identical assortative stable matching in the frictionless case. The interaction between the productive and strategic motives in our model makes it possible to use the data to distinguish empirically between horizontal and vertical preferences. The empirical strategy will depend on data availability, however.

The literature distinguishes two important situations. The first situation is when the contact rates are available. For instance, Hitsch et al (2010) observe the contact rates of both men and women on a dating website, from which one can infer for each type of male and female what their distributions of interest are, $p_x(y)$ and $q_y(x)$. Since these are distributions that sum up to one, in this case, the data contains observations with $2 \times M \times N - M - N$ degrees of freedom. Assuming non-transferable utility, these data allow the researcher to identify the shape of the payoff functions for each type of men and women, ε_{xy} and η_{xy}, which have a total of $2 \times M \times N$ degress of freedom. Our model allows for direct identification of these unobserved preferences up to a constant for each type by using the necessary conditions for equilibrium. Specifically, rearranging equations 3 we obtain:

$$\ln \frac{p_x^*(y)}{p_x^*(y')} = \frac{\varepsilon_{xy}}{\theta_x} q_y^*(x) - \frac{\varepsilon_{xy'}}{\theta_x} q_{y'}^*(x) ,$$

$$\ln \frac{q_y^*(x)}{q_y^*(x')} = \frac{\eta_{xy}}{\theta_y} p_x^*(y) - \frac{\eta_{x'y}}{\theta_y} p_{x'}^*(y) . \tag{4}$$

If the researchers were to restrict attention to cases of equilibrium uniqueness (which is straightforward to test after finding the payoffs), then these equations uniquely identify the best match for each type of male and female, and thus determine whether preferences are horizontal, vertical, or some mix of the two. We were able to routinely recover the correct structure of pref-

erences in Monte-Carlo simulations. Unfortunately we were unable to apply this strategy to the data in Hitsch et al (2010) due to their non-disclosure agreement. As this strategy requires substantial investment in data collection that goes beyond the scope of this paper, we leave the empirical application of this strategy for future research.

The second, more common, situation is when only matching rates are available. For instance, Choo and Siow (2006), observe matching rates for men and women in the U.S. marriage market for specific years, from which one can infer the product $p_x(y) q_y(x)$. In this case, the data contains observations with $M \times N$ degrees of freedom. Of course, there are not enough restrictions in the data to identify payoffs in a non-transferable utility case, but if one were to assume transferable utility with a pre-defined split of the joint payoff between males and females, Φ_{xy}, then the unobserved payoff functions also have $M \times N$ degress of freedom. This makes the payoff functions identifiable in principle.

However, we find that in our model the mapping between the payoff and the matching rate is not necessarily invertible. By that we mean that there may exist matching rate patterns that cannot in principle be generated by our model. Also, we cannot exclude the possibility that some matching rate pattern could be generated by more than one payoff function (although we could not find an example of this in practice).

Given the potential non-invertibility of the mapping between the payoff and the matching rates, our empirical methodology proceeds in three steps. First, we assume that search costs are identical across agents on both sides of the marriage market, $\theta_x = \theta_y = \theta$. This assumption will allow us to identify the ratio of the payoff to search cost, Φ_{xy}/θ, for each pair of types. In addition, we assume that each payoff is split equally between males and

females, i.e. $\epsilon_{xy} = \eta_{xy} = \Phi_{xy}/2$. Second, for any shape of the payoff function, Φ_{xy}, we find all equilibria (if there are more than one) and compute all corresponding equilibrium matching rates implied by the model. Third, we search for a shape of the payoff that maximizes the likelihood function of the data given the predicted matching rates.

Whenever a proposed payoff function produces multiple equilibria, we select the one that fits the observed matching rate best, i.e. has the highest likelihood. Maximization of the likelihood function efficiently minimizes the properly weighted sum of distances between the data and the model's prediction and should lead to consistent estimates. Maximum likelihood estimation of discrete games with multiple equilibria have been reasonably well studied in the literature, e.g. Aguirregabiria and Mira (2007). Here we do not employ any computational tricks since the 3-by-3 case can be computed by brute force in reasonable time. The results of such estimation can be treated as an upper bound on the explanatory power of the model.

In the empirical section, we apply this method to three prominent examples of sorting in the marriage market and find that the model fits the data very well. Despite matching rate data suggesting mostly horizontal preferences, when viewed through the lens of our model with strategic motives, the data are consistent with a vertical preference structure.

4   Empirical Application

To take the model to the data, we use a standard dataset for matching rates in the U.S. marriage market. The data on unmarried males and females and newly married couples comes from IPUMS for the year 2001.[19] For computational transparency we attribute both males and females to three

---

[19]We thank Gayle and Shephard (2015) for kindly sharing the IPUMS data with us.

equally sized bins, which we refer to as low (L), medium (M), and high (H) types. We consider three dimensions along which males and females evaluate each other in the marriage market: income, age, and education. In each case we choose the cutoffs between bins in such a way as to split the whole U.S. population of each gender to equally sized bins.

In the case of age, we restrict our attention only to adults between the ages 21 and 33. To make them as close as possible to equal size, the bins correspond to ages 21-23, 24-27, 28-33. We discard all younger and older people from the analysis because there is a disproportionate amount of unmarried people in these other age categories who only rarely marry. One reason for this may be that a large fraction of them are not searching for a spouse and are thus not participating in the marriage market. To avoid misspecification due to our inability to observe search effort, we exclude them from our analysis. In the case of education, the natural breakdown into three bins is to have people who never attended college, those who are currently in college, and those who have graduated from college. Income is a continuous characteristic, so the three bins correspond to people with low, medium, and high incomes.

For each of the three cases, we estimate the shape of the payoff function using the maximum likelihood methodology described earlier for the case of transferable utility. We assume that all currently unmarried males and females are searching, and the number of matches is proxied by the number of couples that were married in the past 12 months, as indicated by answers to the questionnaire. The dataset contains 93,599 unmarried males, 82,673 unmarried females, and 23,572 newly married couples above the age of 21.

The matching rate for the case of income is presented in the left panel of Figure 3. The estimate of the underlying payoff is shown in the right panel

of the same Figure. A notable property of the payoff is that it shows strong vertical preference. That is, marrying a spouse with a higher income is always better. We find that the matching rate and the payoff have a correlation of 0.72.

The matching rate for the sorting by age is presented in the left panel of Figure 4. Looking at the shape of the matching rate, we would expect to see the horizontal preferences here, with slightly older males looking for slightly older females. However, the shape of the payoff that best explains this sorting pattern is also consistent with vertical preferences. Females have a strong preference for older males independent of their own age. Meanwhile, males are virtually indifferent to the age of their spouse. The highest payoff is produced by males at age 30 marrying females at age 23. The correlation between the matching rate and the payoff is a staggeringly low 0.42.

The matching rate for sorting by education is presented in the left panel of Figure 5. In this case the payoff exhibits segments of both vertical and horizontal preferences. Low educated people and high educated people prefer someone with their same level of education, generating a region of horizontal preferences. However, people with a medium level of education tend to prefer highly educated people, generating a region with vertical preferences. The matching rate and the payoff function have a correlation of 0.52.

A widely used workhorse model of the marriage literature is the model of Choo and Siow (2006). They estimate a static transferable utility model that generates a nonparametric marriage matching function. This model postulates that, in equilibrium, each pair of cohorts of males and females reaches an implicit agreement on the matching rate among themselves; matching (or staying single) is a voluntary decision. In their model, the payoff is recovered as a simple algebraic function of the matching rates and the number of
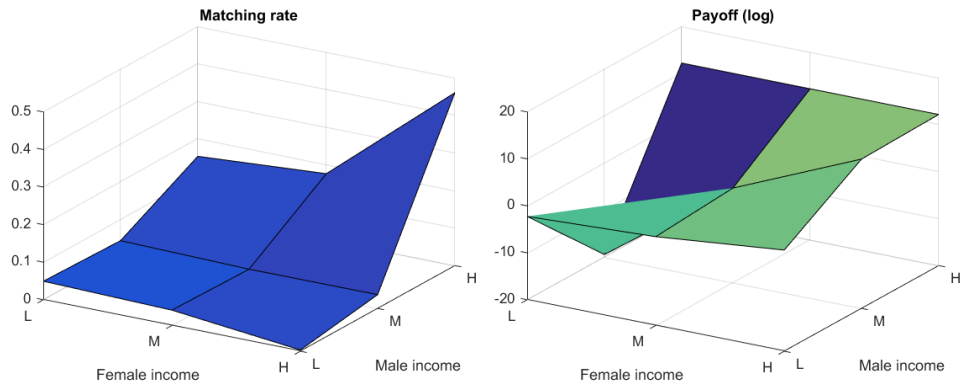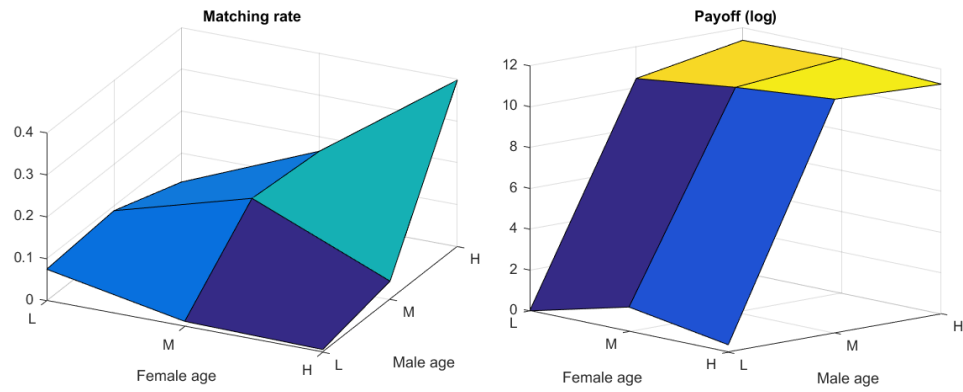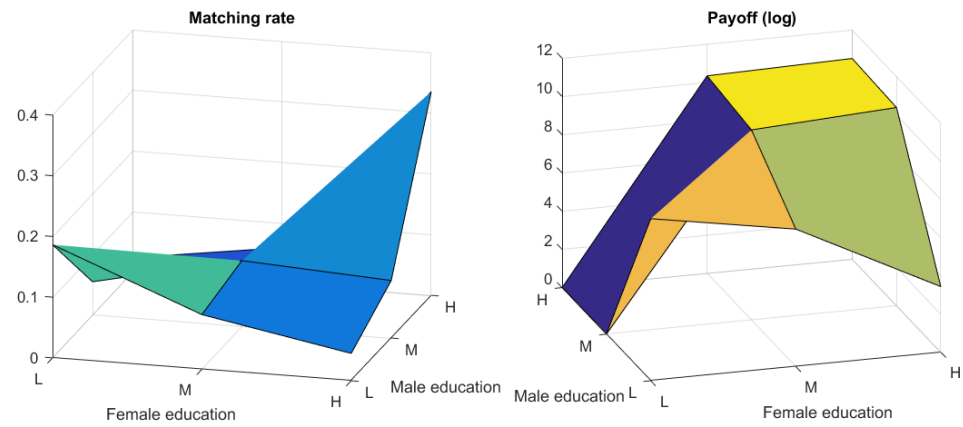
Figure 3: Sorting by income

Figure 4: Sorting by age

Figure 5: Sorting by education

30

people searching. An important feature of the model is that the matching rate depends only on the characteristics of the agents directly involved in the match, but not on the characteristics of other agents present in the marriage market. The strategic motive is absent from their model, so the shape of the matching rate mimics closely the shape of the payoff function. This implies that the distance between the assumptions and implications is minimal: the correlation between the matching rates across pairs of types and the implied values of the payoff are close to one.

We illustrate this feature in Figure 6 where we use the 3-by-3 Monte Carlo simulation from Section 3.2. We plot the correlation between the true underlying payoff and the equilibrium matching rate obtained from our model on the horizontal axis and the correlation between the same matching rate and the corresponding payoff function recovered by the model of Choo and Siow on the vertical axis. We find that in many cases, the shape of the true payoff and of the matching rate descends to 0.4, while the model of Choo and Siow would imply that they have a similar shape with a correlation above 0.75. We color the payoffs consistent the three types of preferences in three different colors. We find that while the correlation depends significantly on the pattern of preferences in our model, in Choo and Siow's model it does not.

The Figure also compares our empirical findings with the Monte Carlo simulation. We find that the three prominent empirical examples that we have considered indeed belong to the range of correlation values commonly generated by payoffs consistent with vertical preferences.
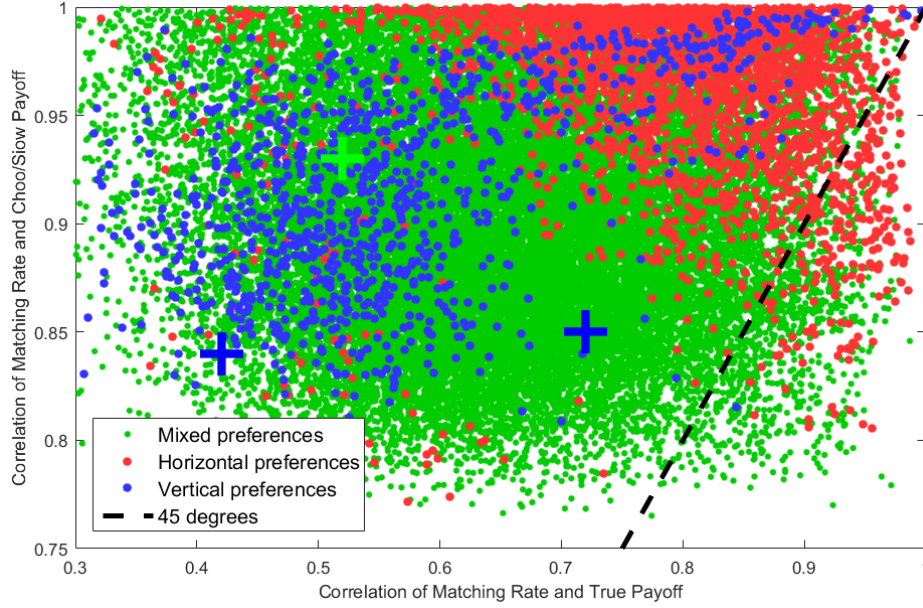
Figure 6: Monte Carlo results and Data

This result emphasizes the importance of considering the effect of strategic motives on the sorting patterns in empirical research. If a researcher looks at the data through the lens of a model with exogenous randomness, that model by construction ignores any strategic considerations that may affect agents' search strategies. As we have shown, strategic considerations can drive a significant wedge between the shape of the preferences and the shape of the observed sorting pattern. Ignoring search frictions that affect the decision of who to meet may thus lead to vastly misleading conclusions regarding the amount of mismatch present in a market and the size of the losses associated with it. In this paper, we have presented and demonstrated the effectiveness of an identification strategy that uses the strategic motives of agents do identify preferences, with the emphasis on distinguishing vertical and horizontal preference structures.

32

5 Final Remarks

In this paper we propose a model of probabilistic choice by agents in a matching market of whom to contact when locating the best partners involves cognitive effort. The model features a productive motive whereby agents target partners that render a higher payoff and a strategic motive which drives agents towards partners who are more likely to show mutual interest. We find that accounting for the interaction of strategic and productive considerations allows the identification of underlying preferences, while ignoring this interaction may result in misleading implications regarding the degree of mismatch and hence the losses associated with it. Understanding who meets whom is crucial for understanding who marries who, and who should marry whom instead.

We applied the model to the U.S. marriage market to demonstrate its relevance, but our model is well suited to study a host of real-life matching markets where agents have limited time and ability to quickly evaluate the relative merits of potential partners. A number of markets ranging from labor markets to education and health care provide examples of markets where equilibrium matches between superior and inferior types are prevalent. Our model can be a useful tool for analyzing these markets.

Furthermore, our model describes markets where the degree of centralization is fairly low. In many two-sided market models a platform acts both as a coordination device and as a mechanism to transfer utility. Our model can be used to study the optimal degree of centralization and the social efficiency of pricing schemes in these markets. We view both the empirical study of matching markets, and the optimal design of centralization in two-sided search environments as exciting areas of future research.

References

[1] **Adachi H. (2003).** "A search model of two-sided matching under non-transferable utility." *Journal of Economic Theory*, 113, pp. 182–198.

[2] **Aguirregabiria, V. and P. Mira. (2007).** "Sequential Estimation of Dynamic Discrete Games." *Econometrica*, 75(1), pp. 1-53, 01.

[3] **Banerjee, A., Duflo, E., Ghatak, M. and J. Lafortune. (2013).** "Marry for What? Caste and Mate Selection in Modern India." *American Economic Journal: Microeconomics*, 5(2), pp. 33-72.

[4] **Becker, G. S. (1973).** "A Theory of Marriage: Part I." *Journal of Political Economy*, 81(4), pp. 813-846.

[5] **Belot, M. and M. Francesconi. (2013).** "Dating Preferences and Meeting Opportunities in Mate Choice Decisions" *Journal of Human Resources*, 48(4), pp. 474-508.

[6] **Burdett, K. and M. G. Coles. (1999).** "Long-Term Partnership Formation: Marriage and Employment." *Economic Journal, Royal Economic Society*, 109(456), pp. F307-F334.

[7] **Chade, H., Eeckhout, J. and L. Smith. (2016).** "Sorting Through Search and Matching Models in Economics." *Journal of Economic Literature*, Forthcoming.

[8] **Cheremukhin, A., Popova A. and A. Tutino. (2015).** "A Theory of Discrete Choice with Information Costs." *Journal of Economic Behavior and Organization*, vol. 113, pp. 34-50.

[9] **Chiaporri, P.-A. and B. Salanie. (2015).** "The Econometrics of Matching Models." *Journal of Economic Literature*, Forthcoming.

[10] **Choo, E. and A. Siow, (2006).** "Who Marries Whom and Why." *Journal of Political Economy*, 114(1), pp. 175-201.

[11] **Eeckhout, J. (1999).** "Bilateral Search and Vertical Heterogeneity." *International Economic Review*, 40(4), pp. 869–887.

[12] **Eeckhout, J. and P. Kircher. (2010).** "Sorting and Decentralized Price Competition." *Econometrica*, 78(2), pp. 539-574.

[13] **Fisman, R., Iyengar, S. S., Kamenica, E. and I. Simonson. (2006).** "Gender Differences in Mate Selection: Evidence from a Speed Dating Experiment." *The Quarterly Journal of Economics*, 121(2), pp. 673-697.

[14] **Galichon, A. and B. Salanié. (2012).** "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models." Working Paper hal-01053710

[15] **Gayle, G.-L. and A. Shephard. (2015).** "Optimal Taxation, Marriage, Home Production, and Family Labor Supply." October 2015, University of Pennsylvania Working Paper.

[16] **Hey, J. D. and C. Orme. (1994).** "Investigating Generalisations of Expected Utility Theory Using Experimental Data." *Econometrica*, 62(6), pp. 1291-1326.

[17] **Herrenbrueck, L., Xia, X., Eastwick, P., Finkel, E. and C. M. Hui. (2016).** "Smart-Dating in Speed-Dating: How a Simple Search

Model Can Explain Matching Decisions." March 2016, Simon Frazer University Working Paper.

[18] **Hitsch, G. J., Hortaçsu, A. and D. Ariely. (2010).** "Matching and Sorting in Online Dating." *American Economic Review*, 100(1), pp. 130–163.

[19] **Lester, B. (2011).** "Information and Prices with Capacity Constraints," *American Economic Review*, 101(4), pp. 1591-1600.

[20] **Matejka F. and A. McKay. (2015).** "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model." *American Economic Review*, 105(1), pp. 272-298.

[21] **Menzio, G. (2007).** "A Theory of Partially Directed Search," *Journal of Political Economy*, 115(5), pp. 748–769.

[22] **Rosen, J.B. (1965).** "Existence and Uniqueness of Equilibrium Points for Concave N-Person Games." *Econometrica*, 33(3), pp. 520-534.

[23] **Shimer, R. and L. Smith. (2000).** "Assortative Matching and Search." *Econometrica*, 68(2), pp. 343-370.

[24] **Shimer, R. (2005).** "The Assignment of Workers to Jobs in an Economy with Coordination Frictions." *Journal of Political Economy*, 113, pp. 996-1025.

[25] **Simsek, A. Ozdaglar, A. and D. Acemoglu. (2007).** "Generalized Poincare-Hopf Theorem for Compact Nonsmooth Regions," *Mathematics of Operations Research*, 32(1), pp. 193-214.

[26] **Smith, G. E., Venkatraman, M. P. and R. R. Dholakia. (1999).** "Diagnosing the Search Cost Effect: Waiting Time and the Moderating

Impact of Prior Category Knowledge," *Journal of Economic Psychology*, 20, pp. 285–314.

[27] **Thomas, C. and J. Cover, (1991)**. "Elements of Information Theory". John Wiley & Sons, Inc.

[28] **Yang, H. (2013).** "Targeted Search and the Long Tail Effect." *The RAND Journal of Economics*, 44(4), pp. 733–756.

NOT FOR PUBLICATION

Appendix A: (In) efficiency of equilibrium

To evaluate the efficiency of the equilibrium, we compare the solution of the decentralized problem to a social planner's solution. We assume that the social planner maximizes the total payoff, which is a utilitarian welfare function. To achieve a social optimum, the planner can choose the strategies of males and females. If there were no search costs, the planner would always choose to match each male with the female that produces the highest output. The socially optimal strategies of males would be infinitely precise.

To study the constrained efficient allocation we impose on the social planner the same costs of search that we place on males and females. Thus, the social planner maximizes the following welfare function:

$$W = \max_{p_x(y), q_y(x)} \sum_{x=1}^{F} \sum_{y=1}^{M} \Phi_{xy} p_x(y) q_y(x) - \sum_{x=1}^{F} c_x(\kappa_x(p_x(y))) - \sum_{y=1}^{M} c_y(\kappa_y(q_y(x)))$$

subject to (1-2) and to the constraints that $p_x(y)$ and $q_y(x)$ are well-defined probability distributions.

Under the assumption of increasing and convex cost functions, the social welfare function is concave in the strategies of males and females. Hence, first-order conditions are necessary and sufficient conditions for a maximum. Rearranging and substituting out Lagrange multipliers, we arrive at the following characterization of the social planner's allocation:

$$p_x^o(y) = \exp\left(\frac{\Phi_{xy} q_y^o(x)}{\theta_x}\right) / \sum_{y'=1}^{M} \exp\left(\frac{\Phi_{xy'} q_{y'}^o(x)}{\theta_x}\right),$$

$$q_y^o(x) = \exp\left(\frac{\Phi_{xy} p_x^o(y)}{\theta_y}\right) / \sum_{x'=1}^{F} \exp\left(\frac{\Phi_{x'y} p_{x'}^o(y)}{\theta_y}\right). \tag{5}$$

The structure of the social planner's solution is very similar to the structure of the decentralized equilibrium given by (3). From a female's perspective, the only difference between the two strategies is that the probability of targeting a male depends on the social gain from a match rather than on her private gain. Notice that the same difference holds from the perspective of a male. Thus, it is socially optimal for both females and males to consider the total payoff, while in the decentralized equilibrium they consider only their private payoffs.

In our model, additional search effort exerted by an individual male or female has a positive externality on the whole matching market. For instance, when a male chooses to increase his search effort, he can better locate his preferable matches. As a consequence, the females he targets will benefit (through an increase in the personal matching rate) and the females he does not target will also be better off as his more targeted strategy will help them exclude him from their search (through a decrease in the personal matching rate). Nevertheless, in this environment agents can not appropriate all the social benefits (the output of a match) they provide to society when increasing their search effort. They only get a fraction of the payoff. This failure of the market to fully compensate both females and males with their social marginal products leads to under-supply of search effort by both sides in the decentralized equilibrium.

Because the social gain is always the sum of private gains, there is no fea-

sible way of splitting the payoff such that it implements the social optimum. When θ is finite and positive, a socially optimal equilibrium has to satisfy the following conditions simultaneously:

$$\varepsilon_{xy} = \Phi_{xy}, \qquad \eta_{xy} = \Phi_{xy}.$$

In the presence of heterogeneity, these optimality conditions can hold in equilibrium only if the total payoff is zero, as private gains have to add up to the total payoff, $\varepsilon_{xy} + \eta_{xy} = \Phi_{xy}$. Therefore, we have just proven the following theorem:

Theorem 3. *The matching equilibrium is socially inefficient for any split of the payoff if all of the following hold:*

1) cost functions are increasing and convex;

2) $\Phi_{xy} > 0$ for some (x, y);

3) $\Phi_{xy} \neq \Phi_{xy'}$ for some x, y and y';

4) $\Phi_{xy} \neq \Phi_{x'y}$ for some y, x and x';

5) $0 < \left. \frac{\partial c_x(\kappa_x)}{\partial \kappa_x} \right|_{p_x^*} = \theta_x < \infty$;

6) $0 < \left. \frac{\partial c_y(\kappa_y)}{\partial \kappa_y} \right|_{q_y^*} = \theta_y < \infty$.

Proof. The proof proceeds in 3 steps.

Step 1. Under the assumption of increasing convex cost functions, both individual payoff functions and the social welfare function are concave in the strategies of males and females. Hence, first-order conditions are necessary and sufficient conditions for a maximum.

Step 2. We denote by CEFOC the first-order conditions of the decentralized equilibrium and by POFOC the first-order conditions of the social planner. In formulae:

$$\text{POFOC}_{p_x(y)}: \qquad \Phi_{xy}\tilde{q}_y(x) - \left. \frac{\partial c_x(\tilde{\kappa}_x)}{\partial \tilde{\kappa}_x} \right|_{\tilde{p}_x(y)} \left(\ln \frac{\tilde{p}_x(y)}{1/M} + 1 \right) - \tilde{\lambda}_x = 0$$

40

$$\text{POFOC}_{q_y(x)}: \qquad \Phi_{xy}\tilde{p}_x(y) - \left.\frac{\partial c_y(\tilde{\kappa}_y)}{\partial \tilde{\kappa}_y}\right|_{\tilde{q}_y(x)}\left(\ln\frac{\tilde{q}_y(x)}{1/F}+1\right) - \tilde{\lambda}_y = 0$$

$$\text{CEFOC}_{p_x(y)}: \qquad \varepsilon_{xy}q_y(x) - \left.\frac{\partial c_x(\kappa_x)}{\partial \kappa_x}\right|_{p_x(y)}\left(\ln\frac{p_x(y)}{1/M}+1\right) - \lambda_x = 0$$

$$\text{CEFOC}_{q_y(x)}: \qquad \eta_{xy}p_x(y) - \left.\frac{\partial c_y(\kappa_y)}{\partial \kappa_y}\right|_{q_y(x)}\left(\ln\frac{q_y(x)}{1/F}+1\right) - \lambda_y = 0$$

For the equilibrium to be socially efficient we need to have the following:

$$\tilde{p}_x(y) = p_x(y) \quad \text{for all } x, y$$

$$\tilde{q}_y(x) = q_y(x) \quad \text{for all } x, y$$

Step 3. By contradiction, imagine that the two conditions above hold. Then, by construction,

$$\left.\frac{\partial c_y(\tilde{\kappa}_y)}{\partial \tilde{\kappa}_y}\right|_{\tilde{q}_y(x)} = \left.\frac{\partial c_y(\kappa_y)}{\partial \kappa_y}\right|_{q_y(x)} = a_y$$

and

$$\left.\frac{\partial c_x(\tilde{\kappa}_x)}{\partial \tilde{\kappa}_x}\right|_{\tilde{p}_x(y)} = \left.\frac{\partial c_x(\kappa_x)}{\partial \kappa_x}\right|_{p_x(y)} = a_x.$$

Denote marginal cost by a_y and a_x respectively. It then follows that:

$$\begin{aligned}
\Phi_{xy}\tilde{p}_x(y) - \tilde{\lambda}_y &= \left.\frac{\partial c_y(\tilde{\kappa}_y)}{\partial \tilde{\kappa}_y}\right|_{\tilde{q}_y(x)}\left(\ln\frac{\tilde{q}_y(x)}{1/M}+1\right)\\
&= \left.\frac{\partial c_y(\kappa_y)}{\partial \kappa_y}\right|_{q_y(x)}\left(\ln\frac{q_y(x)}{1/M}+1\right)\\
&= \eta_{xy}p_x(y) - \lambda_y
\end{aligned}$$

i.e. $\Phi_{xy}\tilde{p}_x(y) - \tilde{\lambda}_y = \eta_{xy}p_x(y) - \lambda_y$ for all x and y. We can use the first-order conditions of males to derive the formulas for λ_y and $\tilde{\lambda}_y$:

$$(i) \qquad F\exp\left(1+\frac{\tilde{\lambda}_y}{a_y}\right) = \sum_{x=1}^{F}\exp\left(\frac{\Phi_{xy}p_x(y)}{a_y}\right)$$

41

$$(ii) \qquad F \exp\left(1 + \frac{\lambda_y}{a_y}\right) = \sum_{x=1}^{F} \exp\left(\frac{\eta_{xy}(x)p_x(y)}{a_y}\right)$$

$$(iii) \qquad \left(\Phi_{xy} - \eta_{xy}\right)p_x(y) = \tilde{\lambda}_y - \lambda_y \qquad \text{for all } x$$

Jointly (i) (ii) and (iii) imply:

$$\frac{\displaystyle\sum_{x'=1}^{F} \exp\left(\frac{\Phi_{x'y}p_{x'}(y)}{a_f}\right)}{\displaystyle\sum_{x'=1}^{F} \exp\left(\frac{\eta_{x'y}p_{x'}(y)}{a_y}\right)} = \frac{\exp\left(\frac{\Phi_{xy}p_x(y)}{a_y}\right)}{\exp\left(\frac{\eta_{xy}p_x(y)}{a_y}\right)} \qquad \text{for all } x$$

Hence,

$$\frac{\exp(\Phi_{xy}p_x(y))}{\exp(\eta_{xy}p_x(y))} = \frac{\exp\left(\Phi_{x'y}p_{x'}(y)\right)}{\exp\left(\eta_{x'y}p_{x'}(y)\right)} \qquad \text{for all } x \text{ and } x'.$$

Therefore, either:

$$\text{a) } \Phi_{xy} = \eta_{xy} \text{ for all } x \text{ or}$$

$$\text{b) } \Phi_{x'y} = \Phi_{x''y} \text{ and } \eta_{x'y} = \eta_{x''y} \text{ for all } x' \text{ and } x'';$$

Similarly from females' first-order conditions it follows that either :

$$\text{c) } \Phi_{xy} = \varepsilon_{xy} \text{ for all } y \text{ or}$$

$$\text{d) } \Phi_{xy'} = \Phi_{xy''} \text{ and } \varepsilon_{x'y} = \varepsilon_{x''y} \text{ for all } y' \text{ and } y''$$

Cases b) and d) have been ruled out by the assumptions of the theorem. Cases a) and b) jointly imply that $\varepsilon_{xy} = \eta_{xy} = \Phi_{xy} = \varepsilon_{xy} + \eta_{xy}$ which leads to a contradiction $\varepsilon_{xy} = \eta_{xy} = \Phi_{xy} = 0$. $\qquad\square$

The first two conditions are self-explanatory; the case when all potential matches yield zero payoffs is a trivial case of no gains from matching. Conditions 5 and 6 state that marginal costs of reducing noise have to be finite and positive in the neighborhood of the equilibrium. When θ is zero,

the best equilibrium of the assignment model is socially optimal. When θ is very high, the random matching outcome is the best possible outcome. For all intermediate values of marginal costs, the decentralized equilibrium is socially inefficient.

Conditions 3 and 4 together require heterogeneity to be two-sided. If heterogeneity is one-sided, i.e. condition 3 or condition 4 is violated, then the allocation of intentions towards the homogeneous side of the market will be uniform. In this case, search becomes one-sided and equilibrium allocations are efficient contingent on the actively searching side appropriating 100 percent of the payoff.[20]

One notable property of the equilibrium is that, by considering only fractions of the total payoff when choosing their strategies, males and females place lower probabilities on pursuing their best matches. This implies that in equilibrium, probability distributions of males and females are more dispersed and the number of matches is lower than is socially optimal.

The inefficiency that arises in the two-sided model can in principle be corrected by a central planner. This can be done by promising both males and females that they will receive the entire payoff of each match and then by collecting lump-sum taxes from both sides of the market to cover the costs of the program. Nevertheless, to do so, the planner himself would need to acquire extensive knowledge about the distribution of the payoffs, which is costly. We leave this point for future research.

---

[20]See Appendix B for a version of the model with one-sided heterogeneity.

Appendix B: One-sided model

Here we consider a one-sided model where females are searching for males who are heterogeneous in type and females face a search cost. We assume that there is no heterogeneity on the female side of the market. As such the probability that a male reciprocates the intentions of a female is given by q_y. The strategy of a female remains $p_x(y)$. Like before a female's cost of searching is given by $c_x(\kappa_x)$. Female x chooses a strategy $p_x(y)$ to maximize her expected income flow:

$$Y_x = \max_{p_x(y)} \sum_{y=1}^{M} \varepsilon_{xy} p_x(y) q_y - c_x(\kappa_x)$$

A female receives her payoff in a match with male y conditional on matching with that male. She also incurs a cost that depends on search effort:

$$\kappa_x = \sum_{y=1}^{M} p_x(y) \ln \frac{p_x(y)}{1/M} \tag{6}$$

where the female's strategy must satisfy $\sum_{y=1}^{M} p_x(y) = 1$ and $p_x(y) \geq 0$ for all y.

Definition 2. *A matching equilibrium of the one-sided matching model is a set of strategies of females, $\{p_x(y)\}_{x=1}^{N}$, which solve their optimization problems.*

Theorem 4. *If the cost functions are non-decreasing and convex, the one-sided matching model has a unique equilibrium.*

Proof. The payoffs of all females are continuous in their strategies. They are also concave in these strategies when cost functions are (weakly) increasing and convex. Hence, each problem has a unique solution. $\square$

When in addition the cost functions are differentiable, it is easy to verify that first-order conditions are necessary and sufficient conditions for equilibrium. Rearranging the first order conditions for the female, we obtain:

$$p_x^* (y) = \exp \left(\frac{\varepsilon_{xy} q_y}{\theta_x} \right) / \sum_{y'=1}^{M} \exp \left(\frac{\varepsilon_{xy'} q_{y'}}{\theta_x} \right). \tag{7}$$

The equilibrium condition (7) has an intuitive interpretation. It predicts that the higher is the female's expected gain from matching with a male, the higher is the probability placed on locating that male. Thus, males are naturally sorted in each female's strategy by probabilities of contacting those males.

Efficiency To study the constrained efficient allocation we impose upon the social planner the same constraints that we place on females. Thus, the social planner maximizes the following welfare function:

$$W = \sum_{x=1}^{F} \sum_{y=1}^{M} \Phi_{xy} p_x (y) q_y - \sum_{x=1}^{F} c_x (\kappa_x)$$

subject to (6) and to the constraint that the $p_x (y)$'s are well-defined probability distributions. Under the assumption of increasing convex cost functions, the social welfare function is concave in the strategies of females. Hence, first-order conditions are sufficient conditions for a maximum. Rearranging we arrive at the following characterization of the social planner's allocation:

$$p_x^o (y) = \exp \left(\frac{\Phi_{xy} q_y}{\theta_x} \right) / \sum_{y'=1}^{M} \exp \left(\frac{\Phi_{xy'} q_y}{\theta_x} \right). \tag{8}$$

Agsain, the structure of the social planner's solution is very similar to the structure of the decentralized equilibrium. The only difference between the centralized and decentralized equilibrium strategies is that the probability of

45

locating a male depends on the social gain from a match rather than on the private gain. Thus, it is socially optimal to consider the whole expected payoff when determining the socially optimal strategies, while in the decentralized equilibrium females only consider their private gains. To decentralize the socially optimal outcome the planner needs to give all of the payoff to the females, $\varepsilon_{xy} = \Phi_{xy}$, effectively assigning them a share of 1. Note that, if the planner could choose the probability that a male reciprocates a female, q_y, he would also set it to 1. When search costs are absent, the equilibrium of the model is socially optimal. When costs are very high, the random matching outcome is the best possible outcome. For all intermediate values of costs, the decentralized equilibrium is constrained efficient contingent on the female receiving the whole output of the match.

Appendix C: Repeated two-sided model

Here we extend the two-sided matching model to a repeated setting. Following Adachi (2003) we assume that at the moment that male y and female x meet, each of them has an additional decision to make. Each agent may choose to form a match and receive the corresponding share of the surplus, or refuse to form a match and wait for a better potential partner in future periods if their continuation value is higher than the utility from matching with the proposed partner. The continuation value is assumed to be simply the expected utility of matching in the future discounted at the rate ρ, which is the patience parameter. In the Adachi model the case $\rho = 1$ represented a frictionless case which implied that agents could wait for their preferred match indefinitely at no time cost to them. Notice, that our one-shot model represents the opposite case of $\rho = 0$.

We denote v_x the continuation value of female x and w_y the continuation value of male y. Each agent chooses her strategy and pays the cost of search before the game starts, and then makes a sequence of draws from the chosen distribution. Matched pairs of agents are replaced by their copies in the search process. The time-0 problems of the agents are like before:

$$Y_x = \sum_{y=1}^{M} EU_x(y)\, q_y(x)\, p_x(y) - \theta_x \left(\sum_{y=1}^{M} p_x(y) \ln \frac{p_x(y)}{1/M} \right) + \lambda_x \left(1 - \sum_{y=1}^{M} p_x(y) \right),$$

$$Y_y = \sum_{x=1}^{F} EU_y(x)\, p_x(y)\, q_y(x) - \theta_y \left(\sum_{x=1}^{F} q_y(x) \ln \frac{q_y(x)}{1/F} \right) + \lambda_y \left(1 - \sum_{x=1}^{F} q_y(x) \right).$$

The continuation values are defined as the solutions to the Bellman programs:

$$v_x = \rho \sum_{y=1}^{M} EU_x(y)\, q_y(x)\, p_x(y) + \rho \left(1 - \sum_{y=1}^{M} q_y(x)\, p_x(y) \right) v_x,$$

$$w_y = \rho \sum_{x=1}^{F} EU_y(x) \, p_x(y) \, q_y(x) + \rho \left(1 - \sum_{x=1}^{F} p_x(y) \, q_y(x) \right) w_y.$$

and the expected utilities from meeting are either equal to match utilities if both partners agree to a match, or to continuation values if they do not reach an agreement:

$$EU_x(y) = v_x + (\eta_{xy} - v_x) \, I \, (\eta_{xy} \geq v_x) \, I \, (\varepsilon_{xy} \geq w_y),$$

$$EU_y(x) = w_y + (\varepsilon_{xy} - w_y) \, I \, (\eta_{xy} \geq v_x) \, I \, (\varepsilon_{xy} \geq w_y).$$

An equilibrium of the model is a set of strategies $\{p_x(y)\}_{x=1}^{F}$, $\{q_y(x)\}_{y=1}^{M}$, reservation values $\{v_x\}_{x=1}^{F}$, $\{w_y\}_{y=1}^{M}$, and expected utilities $\{EU_x(y)\}_{x=1}^{F}$, $\{EU_y(x)\}_{y=1}^{M}$ that jointly solve the problems of the agents and satisfy the system of equations above. Since the maximization problems are well-defined, the first-order conditions are still necessary conditions and must be satisfied in equilibrium. However, because the remaining functions are continuous, but not everywhere differentiable, the model may have multiple equilibria for many different combinations of parameters and it is hard to establish definitive results regarding uniqueness.

So far, this model explicitly postulates non-transferable utility, but it can easily be extended to the case of transferable utility. Specifically, the TU case allows for redistributing the surplus in the cases when joint surplus of the match excees the sum of continuation values of the agents. Therefore, the last two equations are replaced in the TU case by:

$$EU_x(y) = v_x + \left(\eta'_{xy} - v_x \right) I \, (\eta_{xy} + \varepsilon_{xy} \geq v_x + w_y),$$

$$EU_y(x) = w_y + \left(\varepsilon'_{xy} - w_y \right) I \, (\eta_{xy} + \varepsilon_{xy} \geq v_x + w_y).$$

where the utilities adjusted for the payments are defined as:

$$\eta'_{xy} = v_x + \frac{\eta_{xy}}{\Phi_{xy}} \left(\eta_{xy} + \varepsilon_{xy} - v_x - w_y \right),$$

$$\varepsilon'_{xy} = w_y + \frac{\varepsilon_{xy}}{\Phi_{xy}} \left(\eta_{xy} + \varepsilon_{xy} - v_x - w_y \right).$$

Note that in the one-shot model of the main text the TU case and the NTU case are identical as the continuation values are zero. In Figure 7 below we illustrate, using a simple payoff structure that exhibits vertical preferences for 3 males and 3 females, the regions of the parameter space (θ, ρ) in which the equilibrium is non-unique (shaded), as well as the number of pairs of types that are matched in equilibrium with non-zero probability. The case with 3 pairs represents one-to-one matching, while the case with 9 pairs implies that all possible pairings are observed.
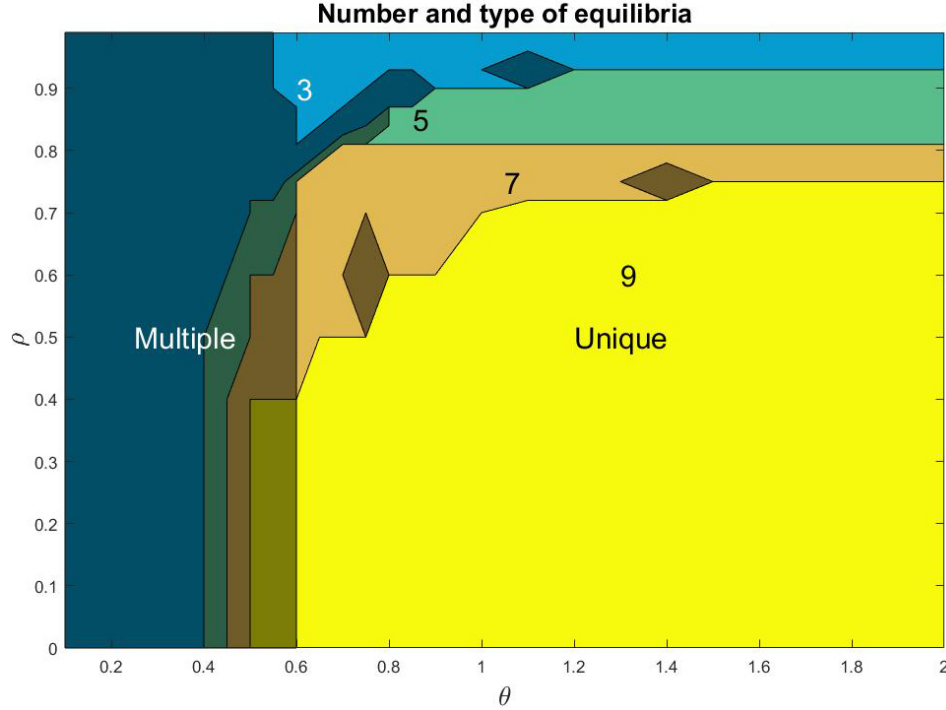

Figure 7: Number and types of equilibria depending on parameters