Targeted Search in Matching Markets

Authors	Anton Cheremukhin, Paulina Restrepo-Echavarria, and Antonella Tutino
Working Paper Number	2014-035F
Revision Date	September 2019
Citable Link	https://doi.org/10.20955/wp.2014.035
Suggested Citation	Cheremukhin, A., Restrepo-Echavarria, P., Tutino, A., 2019; Targeted Search in Matching Markets, Federal Reserve Bank of St. Louis Working Paper 2014-035. URL https://doi.org/10.20955/wp.2014.035

Published In	Journal of Economic Theory
Publisher Link	https://doi.org/10.1016/j.jet.2019.104956

Targeted Search in Matching Markets*

Anton Cheremukhin, Paulina Restrepo-Echavarria, Antonella Tutino

September 20, 2019

Abstract

We propose a parsimonious matching model where a person's choice of whom to meet endogenizes the degree of randomness in matching. The analysis highlights the interaction between a productive motive, driven by the surplus attainable in a match, and a strategic motive, driven by reciprocity of interest of potential matches. We find that the interaction between these two motives differs with preferences—vertical versus horizontal—and that this interaction implies that preferences recovered using our model can look markedly different from those recovered using a model where the degree of randomness is not endogenous. We illustrate these results using data on the U.S. marriage market and show that the model can rationalize the finding of aspirational dating.

JEL: E24, J64, C78, D83.
Keywords: Search, matching, endogenous randomness, preferences.

*Anton Cheremukhin: Federal Reserve Bank of Dallas, 2200 N Pearl St, Dallas TX 75201, chertosha@gmail.com, 214-922-6785. Paulina Restrepo-Echavarria: Federal Reserve Bank of St. Louis, One Federal Reserve Bank Plaza, Broadway and Locust St., St. Louis MO 63166, paulinares@me.com. Antonella Tutino: Federal Reserve Bank of Dallas, 2200 N Pearl St, Dallas TX 75201, tutino.antonella@gmail.com, 214-922-6804.

The authors thank Fernando Alvarez, Gady Barlevy, Hector Chade, Jan Eeckhout, Mike Golosov, Christian Hellwig, Oleg Itskhoki, Boyan Jovanovic, Philipp Kircher, Ricardo Lagos, Rasmus Lentz, Jim Peck, Robert Ulbricht, Pierre-Olivier Weill, and David Wiczer, as well as seminar participants at multiple conferences for discussions and suggestions that helped improve the paper. Authors also thank three anonymous referees, one anonymous associate editor, and editor Alessandro Pavan for comments that improved the paper substantially. All remaining errors are our own. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas, St. Louis or the Federal Reserve System. First draft: January 6, 2012.

1 Introduction

When searching for a match in the dating or marriage market, circumstances can influence whom you meet, but individuals' choices matter as well. As a result, the empirical literature has shown that matching is neither purely random nor perfectly assortative.[1]

In this paper, we propose a parsimonious way to model the choice of whom to meet that endogenizes the degree of randomness in the matching process and show that the model can be used to estimate underlying preferences. Furthermore, we show that the preferences recovered using our model can look markedly different from those recovered using a model where the degree of randomness is not endogenous.

Finally, while our model has a multitude of applications beyond the marriage market, we present one example for which our model is particularly empirically relevant. Bruch and Newman (2018) find that when searching for a match, individuals pursue partners who are "out of their league" in some characteristics, and we show that this is a behavior that derives naturally from our environment.

We blend the stochastic discrete choice literature with the frictionless matching environment of Becker (1973) with two-sided heterogeneity and assume that, on both sides of the market, individual types of agents are characterized by multidimensional attributes. Even though agents know the distribution and their preferences over types, they do not know where to find a particular type. To do so, they decide how much effort they want to exert to locate a particular partner by trading off the cost of search with the payoff they can achieve if successful in finding their desired match.

An agent chooses whom to contact in a probabilistic way, and the strategies chosen are discrete probability distributions over types. Each element of the distribution represents the probability with which an agent will target (i.e., contact) each potential match based on the agent's expected payoff. Exerting more search effort, which results in a higher search cost, allows agents to spot a particular type more accurately. Given the discrete nature of the probability distributions, we model the search cost as proportional to the distance between an uninformed—uniform—strategy, where every

[1] See the survey by Chiapori and Salanie (2015).

type has the same probability of being contacted and the distribution that is chosen by the agent.[2]

The optimal probability distribution representing an agent's contacting strategy balances two motives: the productive and the strategic. The productive motive pushes the agent to pursue the potential match that gives the agent the highest payoff. The strategic motive pushes the agent to pursue the potential match that is more likely to reciprocate interest. Thus, people act strategically not only when deciding whether to form a match or wait for a better option (like in Eeckhout (1999)), but also when choosing whom to contact.

The interaction between the productive and the strategic motives determines the meeting rates in the model. The relative strength of the two motives depends on the search cost. When exerting effort to find the best partner is not very costly, it is easy for agents to locate their preferred types accurately, and reciprocity of interest is the paramount determinant of who meets whom: The strategic motive dominates.

When exerting search effort is costly, agents will not be able to locate their preferred types with accuracy, so the likelihood of contacting someone else increases. In this case, payoffs become the driving force behind who meets whom. In the unique equilibrium, every agent's strategy is to target the partner that would yield the highest payoff: The productive motive dominates the strategic motive.

The equilibrium is generally inefficient. Two externalities prevent the competitive equilibrium from achieving the social optimum. The first is a positive externality. If an agent increases her search effort, not only does she increase the probability of finding a match, but she also generates an incentive for others to increase their search effort, which can increase the quality of matches. Since the agent is not compensated for this additional effort, she fails to internalize this gain. The second externality is negative. When the productive motive drives optimal individual strategies, there is more congestion because competition increases for types targeted with higher intensity, undermining matching probability. Although the equilibrium is generally inefficient,

[2] We borrow our cost specification from the literature on discrete choice under information frictions. See Cheremukhin et al. (2015) and Matejka and McKay (2015).

it is possible for the two externalities to offset each other to achieve efficiency in the competitive equilibrium.

The model can be used for empirical identification of agents' preferences. In particular, it can identify whether agents' preferences are vertical—attraction is based on a commonly agreed upon ranking—or horizontal—people are attracted to agents with similar characteristics.[3] The existing literature finds it hard to distinguish between these cases empirically.[4] 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 motives pull agents in the same direction, as similar types both get the highest payoff from each other and their interests are also more likely to be mutual. In this case, a stochastic version of assortative matching is preserved in equilibrium and the shapes of the observed matching rate and the underlying payoff function are similar.

However, in the case of vertical preferences, there is common agreement on the ranking of agents and everybody would like to chase a particular type but there would be a lower probability of reciprocation. As a result, in this case the productive and the strategic motives pull in opposite directions. The productive motive will encourage agents to target their preferred type, and the strategic motive will encourage them to diversify to increase the chance of forming a match. This case gives rise to a novel equilibrium pattern that resembles neither positive nor negative assortative matching. We call it a mixing equilibrium.[5] This type of equilibrium rationalizes the behavior observed by Bruch and Newman (2018) in online dating where so-called aspirational dating, or "reaching up the desirability ladder," is documented. Such behavior—and sorting in equilibrium—implies that there can be a large difference between the shape of the observed matching rate and the underlying payoff function (preferences).

Our model is not the first one where the shape of the matching rate differs from the shape of the underlying payoff function. This difference arises in models that have a

[3]See Banerjee et al. (2013) and references therein.
[4]See Hitsch et al. (2010) and related studies.
[5]Our taxonomy of equilibria in this case follows that of Burdett and Coles (1999).

strategic motive (see Eeckhout (1999), Shimer and Smith (2000), Chade (2001), Adachi (2003), and Eeckhout and Kircher (2010) among many others). However, our model is the first that can rationalize specific patterns of strategic behavior such as reaching up the desirability ladder. What makes our model different is the endogenous source of randomness in the matching equilibrium, unlike in the existing literature where randomness is postulated.[6]

We provide conditions for existence, uniqueness, and efficiency of equilibrium as well as a characterization of equilibrium sorting. Our results on equilibrium uniqueness allow us to develop a methodology for recovering agents' preferences using only data on aggregate matching rates.

We find that the model does a very good job rationalizing the observed aggregate matching rates in the U.S. marriage market based on education, race, income, and age. Furthermore, we show that due to the presence of the strategic motive—for which relative importance is determined endogenously—the degree of horizontality of the recovered payoff function can be different from that recovered using models in which there is no strategic component or the relative importance of this motive is built in ad hoc. Finally, the resulting strategies—who targets whom—are consistent with the desirability estimates of Bruch and Newman (2018) derived from online dating data.

The paper proceeds as follows: Section 2 describes the model and Section 3 derives the theoretical results. We develop an empirical methodology for recovering preferences and apply it to the U.S. marriage market data in Section 4. Section 5 states some final remarks.

Related literature

Our paper effectively blends two sources of randomness used in the 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

[6]Existing models cannot explain reaching up the desirability ladder because they either postulate random search, where contacts are uniformly drawn from the pool of searchers, or directed search, where an assortative matching outcome ensures that agents match partners of similar rank.

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' choices 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 choices, rather than by some unobserved characteristics with fixed distributions.

The search and matching literature has seen multiple attempts to produce intermediate degrees of randomness with which agents meet their best matches. In particular, Menzio (2007) and Lester (2011) nest directed search and random matching to generate outcomes with an intermediate degree of randomness.[7] Our paper produces equilibrium outcomes in between uniform random matching and the frictionless assignment, endogenously, without nesting these two frameworks.

Also note that although the directed search literature, such as Eeckhout and Kircher (2010) and Shimer (2005), technically involves a choice of whom to meet, the choice is degenerate—directed by signals from the other side. See Chade, Eeckhout, and Smith (2017) for a thorough 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 simultaneously search 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.

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

We model this by assuming that each agent chooses a discrete probability distribution over types. Each element of 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 (or agent) is associated with a higher cost, as the agent needs to exert more effort to locate a particular person more accurately.

2.1 Population and agents

The economy contains a large, finite number of individual agents: females whose types are indexed by $x \in \{1, ..., F\}$ and males whose types are indexed by $y \in \{1, ..., M\}$. We denote by μ_x the number of females of type x and by μ_y the number of males of type y. Identities of females of type x are indexed by $i \in \{1, ..., \mu_x\}$, and identities of males of type y are indexed by $j \in \{1, ..., \mu_y\}$. We think of females and males characterized by a multidimensional set of attributes (e.g., income, age, education, hobbies). For example, the profile of a female of type x' that includes her education, income, and race, among other characteristics, makes her distinct from another female of type x'' with, e.g., the same race but different education and income levels. Note that types x and y are unranked indices that aggregate all attributes.

2.2 Actions

When seeking to form a match, both females and males are aware of the number of agents of each type and the characteristics of their preferred types on the other side of the market. They face a noisy search process where they are uncertain about how to locate their preferred partner. In this environment, each agent's action is a probability distribution over agents on the other side of the market. Since the number of potential partners is finite, the strategy of each agent is a discrete probability distribution.

Let $p_{x,i}(y, j)$ be the probability that a female i of type x targets a male j of type y. Similarly, we denote by $q_{y,j}(x, i)$ the probability that a male j of type y targets a female i of type x.

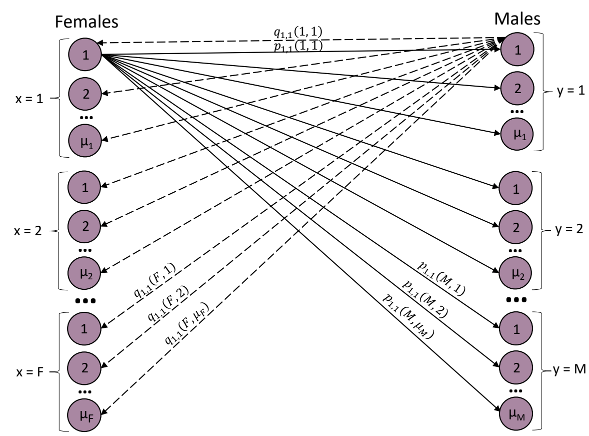

Figure 1: Strategies of Males and Females

The set of actions $s \in S$ is given by the cartesian product of the sets of strategies of females $p_{x,i}(y,j) \in S_{x,i}$ and males $q_{y,j}(x,i) \in S_{y,j}$, where

$$S_{x,i} = \left\{ p_{x,i}(y,j) \in R^{\left(\sum_{y=1}^{M} \mu_y\right)} : p_{x,i}(y,j) \geq 0, \sum_{y=1}^{M} \sum_{j=1}^{\mu_y} p_{x,i}(y,j) \leq 1 \right\},$$

$$S_{y,j} = \left\{ q_{y,j}(x,i) \in R^{\left(\sum_{x=1}^{F} \mu_x\right)} : q_{y,j}(x,i) \geq 0, \sum_{x=1}^{F} \sum_{i=1}^{\mu_x} q_{y,j}(x,i) \leq 1 \right\}.$$

Figure 1 illustrates the strategies of males and females. Consider a female $i = 1$ of type $x = 1$. The solid arrows show the probability $p_{1,1}(y,j)$ she assigns to targeting a male j of each type y. Similarly, dashed arrows show the probability $q_{1,1}(x,i)$ that a male $j = 1$ of type $y = 1$ assigns to targeting a female i of type x. Once these are selected, both males and females make one draw from their respective distributions to determine which individual they will contact.

We make the following assumption on the strategies of individual agents:

A.1 Individual strategies are non-cooperative within and across types.

Assumption A.1 states that an individual agent chooses her strategy taking as given the actions of the other agents on both her side and the opposite side of the market.[8]

[8] Think, for example, about a dating website. When an agent joins the site, there are many members of the different types, and the agent joining thinks that her strategy of searching and contacting potential matches has no effect on the strategies of fellow female or male members.

2.3 Payoffs

A match between any female of type x and any male of type y generates a payoff (surplus) Φ_{xy}. Note that we do not place any restrictions on the shape of the payoff function. 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 $\Phi_{xy} = \varepsilon_{xy} + \eta_{xy}$. The payoff and the split generated by any potential (x, y) match are exogenous and known ex-ante and are independent of their identities i and j, respectively.[9]

Agents form a match if they meet according to the meeting technology (see Subsection 2.5) and each agent (weakly) benefits from forming a match; i.e., each agent's payoff is non-negative. Since a negative payoff corresponds to absence of a match, we make the following assumption on the payoffs:

A.2 The payoffs are non-negative:

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

2.4 Costs of search

When searching for a partner, both types of agents face a noisy search process. In particular, we postulate that individuals cannot perfectly locate potential candidates with preferred characteristics on the other side of the market. Reducing the noise to locate a potential partner more accurately is costly: it involves a careful analysis of the profiles of potential matches, with considerable effort in sorting through the multifaceted attributes of each candidate. When seeking to form a match, agents rationally weigh costs and benefits of targeting the type characteristics that result in a suitable match.

A female rationally chooses her strategy $p_{x,i}(y, j)$ by balancing the costs and benefits of targeting a potential match. A strategy $p_{x,i}(y, j)$ that is more concentrated on a particular male j of type y affords her a higher probability to be matched with her

[9]As we discuss at the beginning of Section 3, the fact that fundamentals such as payoffs and costs are identity-independent will allow us to later drop the indexes i and j.

preferred male. However, it requires more effort to sort through profiles of all the males in the market to locate her desired match and exclude the others.

We assume that agents enter the search process with a uniform prior of whom to target, $\tilde{p}_{x,i}(y,j) = 1/\sum_{y=1}^{M}\mu_y$ and $\tilde{q}_{y,j}(x,i) = 1/\sum_{x=1}^{F}\mu_x$. Choosing a more targeted strategy implies a larger distance between the chosen strategy and the uniform prior and is associated with a higher search effort. A natural way to introduce this feature into our model is the Kullback-Leibler divergence (relative entropy),[10] which provides a convenient way of quantifying the distance between any two distributions, including discrete distributions as in our model. We assume that the search effort of female i of type x is defined as follows:

$$\kappa_{x,i} = \sum_{y=1}^{M}\sum_{j=1}^{\mu_y} p_{x,i}(y,j)\ln\frac{p_{x,i}(y,j)}{\tilde{p}_{x,i}(y,j)}. \tag{1}$$

We assume that the search costs $c_{x,i}(\kappa_{x,i})$ are a function of the search effort $\kappa_{x,i}$. Note that $\kappa_{x,i}$ is increasing in the distance between a uniform distribution over males and the chosen strategy, $p_{x,i}(y,j)$. If an agent does not want to exert any search effort, she can choose a uniform distribution over types and meet males randomly. As she chooses a more targeted strategy, the distance between the uniform distribution and her strategy $p_{x,i}(y,j)$ grows, increasing search effort $\kappa_{x,i}$ and the overall cost of search. By increasing search effort, agents bring down uncertainty about locating a prospective match, which allows them to target their better matches more accurately.

Likewise, a male's cost of search $c_{y,j}(\kappa_{y,j})$ is a function of the search effort defined as

$$\kappa_{y,j} = \sum_{x=1}^{F}\sum_{i=1}^{\mu_x} q_{y,j}(x,i)\ln\frac{q_{y,j}(x,i)}{\tilde{q}_{y,j}(x,i)}. \tag{2}$$

Furthermore, we assume the following:

[10] In the model of information frictions used in the rational inattention literature, κ_x represents the relative entropy between a uniform prior and the posterior strategy. 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.

A.3 The search costs of agents $c_x(\kappa)$ and $c_y(\kappa)$ are strictly increasing, twice continuously differentiable and (weakly) convex functions of search effort.

As a special case, consider a linear cost of search. Then, the total costs of search for a female i of type x are given by $c_{x,i} = \theta_x \kappa_{x,i}$ and for a male j of type y by $c_{y,j} = \theta_y \kappa_{y,j}$, where $\theta_x \geq 0$ and $\theta_y \geq 0$ are the marginal costs of search.

2.5 Meeting technology

The meeting rate depends on the strategies of each agent, $p_{x,i}(y,j)$ and $q_{y,j}(x,i)$, and a congestion function $\phi\left(\{p_{x,i'}(y,j'), q_{y,j'}(x,i')\}_{i',j'}, \mu_x, \mu_y\right)$, which depends in some general way on the strategies of all other agents as well as the number of agents of each type. Given this, the total number of matches formed between females of type x and males of type y is given by

$$\mathfrak{M}_{x,y} = \sum_{i=1}^{\mu_x} \sum_{j=1}^{\mu_y} p_{x,i}(y,j)\, q_{y,j}(x,i)\, \phi\left(\{p_{x,i'}(y,j'), q_{y,j'}(x,i')\}_{i' \in \{1,\ldots,\mu_x\}, j' \in \{1,\ldots,\mu_y\}}, \mu_x, \mu_y\right).$$

We make the following assumption on the structure of the congestion function:

A.4 The congestion function is twice continuously differentiable in each p and q.

We introduce this congestion function following Shimer and Smith (2001) and Mortensen (1982), who assume a linear search technology. Note that if $\phi(\ldots) = 1$, then a match takes place if and only if there is mutual coincidence of interests; i.e., both agents draw each other out of their respective distribution of interests. By introducing this congestion function we are allowing for matches to depend in some general way on both an agent's search intensity for a specific agent (p and q) and on the number of agents taking part.[11]

Note that when setting up the congestion function we implicitly assume that there are no direct inter-type congestion externalities. However, our model still features strong indirect equilibrium interactions between the strategies of agents that work akin

[11] Note that here search intensity refers to how concentrated the distribution of interests of an agent is. A higher search intensity results in assigning higher probability to one or several agents within an agent's distribution of interests.

to inter-type congestion by attracting or deterring agents. We come back to this point in Subsection 3.1.

Given the above, the meeting rate faced by a male of type y conditional on targeting a female of type x takes on the form

$$P_{x,i}\left(y,j\right) = p_{x,i}\left(y,j\right)\phi\left(\left\{p_{x,i'}\left(y,j'\right), q_{y,j'}\left(x,i'\right)\right\}_{i',j'}, \mu_x, \mu_y\right),$$

and, similarly, the meeting rate faced by a female of type x conditional on targeting a male of type y is

$$Q_{y,j}\left(x,i\right) = q_{y,j}\left(x,i\right)\phi\left(\left\{p_{x,i'}\left(y,j'\right), q_{y,j'}\left(x,i'\right)\right\}_{i',j'}, \mu_x, \mu_y\right).$$

Since each individual agent is "small" compared with the population of agents of his type, we assume the following:

A.5 Agents take the meeting rates they face as given, disregarding the dependence of the congestion function on agents' own search intensities.[12]

2.6 Equilibrium

Both males and females maximize the expected value of their payoffs net of the search costs. For a female i of type x, the problem is

$$Y_{x,i} = \max_{p_{x,i}(y,j)\in S_{x,i}} \sum_{y=1}^{M}\sum_{j=1}^{\mu_y} \varepsilon_{xy} Q_{y,j}\left(x,i\right) p_{x,i}\left(y,j\right) - c_x\left(\kappa_{x,i}\left(p_{x,i}\left(y,j\right)\right)\right) \tag{3}$$

for all $i \in \{1,...,\mu_x\}$, $x \in \{1,...,F\}$, where the meeting rates $Q_{y,j}\left(x,i\right)$ are taken as given.

Likewise, a male j of type y solves

$$Y_{y,j} = \max_{q_{y,j}(x,i)\in S_{y,j}} \sum_{x=1}^{F}\sum_{i=1}^{\mu_x} \eta_{xy} P_{x,i}\left(y,j\right) q_{y,j}\left(x,i\right) - c_y\left(\kappa_{y,j}\left(q_{y,j}\left(x,i\right)\right)\right) \tag{4}$$

for all $j \in \{1,...,\mu_y\}$, $y \in \{1,...,M\}$, where the meeting rates $P_{x,i}\left(y,j\right)$ are taken as given. The above expressions allow for a precise definition of the equilibrium:

[12] Intuitively, going back to the dating website example, A.5 states that individual members take the technology behind the interface of the matching website as given and they do not consider how their own strategies may affect the meeting rates they face.

Definition 1. *A* matching equilibrium *is a set of admissible strategies for females,* $\{p_{x,i}(y,j) \in S_{x,i}\}_{x\in\{1,\ldots,F\},i\in\{1,\ldots,\mu_x\}}$*, males,* $\{q_{y,j}(x,i) \in S_{y,j}\}_{y\in\{1,\ldots,M\},j\in\{1,\ldots,\mu_y\}}$*, and meeting rates* $P_{x,i}(y,j)$ *and* $Q_{y,j}(x,i)$*, such that the strategies solve the problems in (3) and in (4) for each individual male and female given the meeting rates, which are consistent with the strategies of the agents.*

3 Results

Note that, given the assumptions made about the surplus and the meeting technology, all agents of the same type are identical ex-ante, and as long as their problems are well defined, we assume that identical agents choose identical actions.[13] Therefore, $p_{x,i}(y,j) = p_{x,i'}(y,j')$ and $q_{y,j}(x,i) = q_{y,j'}(x,i')$ for all $i \neq i'$ and $j \neq j'$. This allows us to drop the index on individual identities i and j and rewrite the problems of the agents as follows:

$$Y_x = \max_{p_x(y)\in S_y} \sum_{y=1}^{M} \mu_y \varepsilon_{xy} Q_y(x) p_x(y) - c_x \left(\sum_{y=1}^{M} \mu_y p_x(y) \ln \left(p_x(y) \sum_{y=1}^{M} \mu_y \right) \right), \qquad (5)$$

$$Y_y = \max_{q_y(x)\in S_y} \sum_{x=1}^{F} \mu_x \eta_{xy} P_x(y) q_y(x) - c_y \left(\sum_{x=1}^{F} \mu_x q_y(x) \ln \left(q_y(x) \sum_{x=1}^{F} \mu_x \right) \right), \qquad (6)$$

where meeting rates are defined as $P_x(y) = p_x(y)\phi_{xy}$, $Q_y(x) = q_y(x)\phi_{xy}$, with $\phi_{xy} = \phi(p_x(y), q_y(x), \mu_x, \mu_y)$ to simplify notation, and admissible strategies satisfy

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

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

[13]This assumption is similar in spirit to the anonymity assumption employed in Shimer (2005). However, in our model this assumption becomes a prediction whenever the equilibrium is unique: identical agents will rationally choose identical strategies. In cases when the congestion function produces multiple equilibria some of those might include asymmetric strategies chosen by identical agents. In those cases, by assumption we exclude asymmetric equilibria from consideration.

3.1 Characterization of equilibrium

We can re-write the objective functions of the agents introducing the linear constraints on strategies via Lagrange multipliers (λ_x and λ_y). Then the first-order conditions for optimality are

$$\frac{\partial Y_x}{\partial p_x(y)} = \varepsilon_{xy} Q_y(x) - \frac{\partial c_x}{\partial \kappa_x}\left(\ln p_x(y) \sum_{y=1}^{M} \mu_y + 1\right) - \lambda_x = 0, \tag{7}$$

$$\frac{\partial Y_y}{\partial q_y(x)} = \eta_{xy} P_x(y) - \frac{\partial c_y}{\partial \kappa_y}\left(\ln q_y(x) \sum_{x=1}^{F} \mu_x + 1\right) - \lambda_y = 0. \tag{8}$$

Since the objective functions of agents are twice continuously differentiable and concave in their own strategies,[14] first-order conditions are necessary and sufficient conditions for equilibrium. Rearranging and substituting out Lagrange multipliers, we obtain the following proposition:

Proposition 1. *Under assumptions A.1-A.5, a matching equilibrium satisfies*

$$p_x^*(y) = exp\left(\frac{\varepsilon_{xy} q_y^*(x)\,\phi_{xy}^*}{\partial c_x/\partial \kappa_x|_{p_x^*(y)}}\right) / \sum_{y'=1}^{M} \mu_{y'} exp\left(\frac{\varepsilon_{xy'} q_{y'}^*(x)\,\phi_{xy'}^*}{\partial c_x/\partial \kappa_x|_{p_x^*(y')}}\right), \tag{9}$$

$$q_y^*(x) = exp\left(\frac{\eta_{xy} p_x^*(y)\,\phi_{xy}^*}{\partial c_y/\partial \kappa_y|_{q_y^*(x)}}\right) / \sum_{x'=1}^{F} \mu_{x'} exp\left(\frac{\eta_{x'y} p_{x'}^*(y)\,\phi_{x'y}^*}{\partial c_y/\partial \kappa_y|_{q_y^*(x')}}\right). \tag{10}$$

Each agent of either type optimally chooses whom to target based on two motives: the productive and the strategic. The productive motive leads the agent to seek out the most desirable type on the opposite side of the market based on the payoff (given by ε_{xy} and η_{xy}). The strategic motive, on the other hand, leads the agent to go after someone who is more likely to reciprocate interest (given by q_y^* and p_x^*).

The congestion function, ϕ_{xy}^*, can scale up or down the probability of forming a match between individuals of type x and y and, as such, will affect optimal strategies (we come back to this point in Subsection 3.5).

[14]Note that weak convexity in A.3 is enough since the search effort (κ) is strictly convex in strategies, and that, A.3 ensures that all equilibrium strategies are strictly positive.

The fact that agents cannot locate their preferred match with accuracy mitigates the strategic motive by increasing uncertainty about attributes of potential matches; this, in turn, may increase the probability of a type with less-than-desirable attributes seeking out a partner with more desirable attributes if the surplus of the match is high enough. As a result, higher search costs boost the productive motive.

Even though in the model there are no direct inter-type congestion externalities, it still features strong indirect equilibrium interactions between the strategies of agents. Because of the strategic motive, if a female x knows that a male y places a high probability on her, the best thing for her to do is to reciprocate by also placing a high probability on y. This will in turn affect the probability a female x' places on y; she will make it lower.

3.2 Existence

The equilibrium of the matching model can be interpreted as a pure-strategy Nash equilibrium of a strategic form game. The following assumption and theorem establish conditions under which a matching equilibrium exists:

A.6 $\phi_{xy} + q_x \frac{\partial \phi_{xy}}{\partial q_y} \geq 0$ and $\phi_{xy} + p_x \frac{\partial \phi_{xy}}{\partial p_x} \geq 0$ for all admissible $p_x(y), q_y(x)$ for all x, y.

This assumption requires that the total matching rate $\mathfrak{M}_{x,y} = \mu_x \mu_y p_x(y) q_y(x) \phi_{xy}$ is non-decreasing in each of the strategies $p_x(y)$ and $q_y(x)$. In other words, it requires that as agents exert more search effort—or increase their search intensity—the matching rate increases.

Theorem 1. *Under assumptions A.1- A.6, a matching equilibrium exists.*

Proof. Since the strategy space is a simplex and, hence, non-empty, convex, and compact set, sufficient conditions for existence of the equilibrium require us to check whether the payoff functions are supermodular on the whole strategy space as in Tarski (1955). Supermodularity can be proven by showing non-negativity of the off-diagonal elements of the Hessian matrix.

Let $J_{xy} = \begin{bmatrix} \frac{\partial Y_y}{\partial q_y} & \frac{\partial Y_x}{\partial p_x} \end{bmatrix}$ be the Jacobian matrix collecting the set of first-order conditions for all $y \in \{1, ..., M\}$ and all $x \in \{1, ..., F\}$ and let H_{xy} be the corresponding

Hessian matrix. To derive the Hessian matrix, note that under A.1, strategies of each individual agent are non-cooperative, i.e., independent of the strategies of opposite types as well as the strategies of the other agents of their own type. Note also that we have assumed no direct inter-type congestion externalities. These assumptions produce a Hessian matrix with a block-diagonal structure, which greatly simplifies the analysis.

For illustrative purposes, suppose that there are four players (two female and two male types).[15] Suitably rearranging the order of strategies for females and males, the Hessian of this eight-action game can be written as

$$H_{xy} = \begin{bmatrix} \frac{\partial^2 Y_{x1}}{\partial p_{11}\partial p_{11}} & \frac{\partial^2 Y_{x1}}{\partial p_{11}\partial q_{11}} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\partial^2 Y_{y1}}{\partial q_{11}\partial p_{11}} & \frac{\partial^2 Y_{y1}}{\partial q_{11}\partial q_{11}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial^2 Y_{x1}}{\partial p_{12}\partial p_{12}} & \frac{\partial^2 Y_{x1}}{\partial p_{12}\partial q_{12}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial^2 Y_{y2}}{\partial q_{21}\partial p_{12}} & \frac{\partial^2 Y_{y2}}{\partial q_{21}\partial q_{21}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\partial^2 Y_{x2}}{\partial p_{21}\partial p_{21}} & \frac{\partial^2 Y_{x2}}{\partial p_{21}\partial q_{12}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\partial^2 Y_{y1}}{\partial q_{12}\partial p_{21}} & \frac{\partial^2 Y_{y1}}{\partial q_{12}\partial q_{12}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\partial^2 Y_{x2}}{\partial p_{22}\partial p_{22}} & \frac{\partial^2 Y_{x2}}{\partial p_{22}\partial q_{22}} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\partial^2 Y_{y2}}{\partial q_{22}\partial p_{22}} & \frac{\partial^2 Y_{y2}}{\partial q_{22}\partial q_{22}} \end{bmatrix}.$$

From this structure, it is clear that for the general case with $y \in \{1, ..., M\}$ and $x \in \{1, ..., F\}$, the off-diagonal elements of H_{xy} are non-negative if, simultaneously,

$$\frac{\partial^2 Y_x}{\partial p_x(y)\,\partial q_y(x)} = \varepsilon_{xy}\left(\phi_{xy} + q_y(x)\,\frac{\partial \phi_{xy}}{\partial q_y(x)}\right) \geq 0,$$

and

$$\frac{\partial^2 Y_y}{\partial q_y(x)\,\partial p_x(y)} = \eta_{xy}\left(\phi_{xy} + p_x(y)\,\frac{\partial \phi_{xy}}{\partial p_x(y)}\right) \geq 0,$$

which requires non-negativity of payoffs as guaranteed by assumption A.2 and that each term in brackets is non-negative, guaranteed by A.6. □

3.3 Uniqueness

For the general proof of uniqueness, we reuse the Jacobian and the Hessian from the previous subsection. Following Rosen (1965) and Gale and Nikaido (1965), if the admissible strategy space S is a convex, non-empty, compact set and the symmetric Hessian

[15] While the extension to the case of M male types and F female types is straightforward, the notation for the general case is cumbersome.

$H_{xy} + H_{xy}^T$ is negative definite for all admissible strategies, then the payoff functions are diagonally strictly concave. We can then use the result that, if the constraints are concave functions, if there exist interior points of the strategy space where the constraints are non-binding, and if the payoff functions are diagonally strictly concave for all admissible strategies, then the game has a unique pure-strategy Nash equilibrium.

Theorem 2. *Under assumptions A.1-A.6 and if the following two conditions hold for all x and y, the matching equilibrium is unique:*

a. $\left.\frac{\partial c_x}{\partial \kappa_x}\right|_{p_x^*(y)} \frac{1}{p_x^*(y)} + \left.\frac{\partial^2 c_x}{\partial \kappa_x \partial \kappa_x}\right|_{p_x^*(y)} \left(\ln p_x^*(y) \sum_{y=1}^{M} \mu_y + 1\right)^2$
$> \varepsilon_{xy} \left(\phi_{xy}^* + q_y^*(x) \frac{\partial \phi_{xy}^*}{\partial q_y(x)} + q_y^*(x) \frac{\partial \phi_{xy}^*}{\partial p_x(y)}\right)$;

b. $\left.\frac{\partial c_y}{\partial \kappa_y}\right|_{q_y^*(x)} \frac{1}{q_y^*(x)} + \left.\frac{\partial^2 c_y}{\partial \kappa_y \partial \kappa_y}\right|_{q_y^*(x)} \left(\ln q_y^*(x) \sum_{x=1}^{F} \mu_x + 1\right)^2$
$> \eta_{xy} \left(\phi_{xy}^* + p_x^*(y) \frac{\partial \phi_{xy}^*}{\partial p_x(y)} + p_x^*(y) \frac{\partial \phi_{xy}^*}{\partial q_y(x)}\right)$.

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 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 must all be non-positive, consistent with concavity of the payoffs:

$$\frac{\partial^2 Y_x}{\partial p_x \partial p_x} = \varepsilon_{xy} q_y(x) \frac{\partial \phi_{xy}}{\partial p_x(y)} - \frac{\partial c_x}{\partial \kappa_x} \frac{1}{p_x(y)} - \frac{\partial^2 c_x}{\partial \kappa_x \partial \kappa_x} \left(\ln p_x(y) \sum_{y=1}^{M} \mu_y + 1\right)^2 \leq 0, \quad (11)$$

$$\frac{\partial^2 Y_y}{\partial q_y \partial q_y} = \eta_{xy} p_x(y) \frac{\partial \phi_{xy}}{\partial q_y(x)} - \frac{\partial c_y}{\partial \kappa_y} \frac{1}{q_y(x)} - \frac{\partial^2 c_y}{\partial \kappa_y \partial \kappa_y} \left(\ln q_y(x) \sum_{x=1}^{F} \mu_x + 1\right)^2 \leq 0. \quad (12)$$

The off-diagonal elements are all non-negative, as we have assumed for equilibrium existence. The remaining cross-derivatives are all zero. To guarantee that the Hessian is negative definite (a stronger condition), 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 the sum of off-diagonal elements, which in turn guarantees that the Hessian of the game is negative definite. When the cost functions are linear (see Section 2.4), these conditions simplify to

$$\theta_x > \varepsilon_{xy} p_x(y)\left(\phi_{xy} + q_y(x)\frac{\partial \phi_{xy}}{\partial q_y(x)} + q_y(x)\frac{\partial \phi_{xy}}{\partial p_x(y)}\right),$$
$$\theta_y > \eta_{xy} q_y(x)\left(\phi_{xy} + p_x(y)\frac{\partial \phi_{xy}}{\partial p_x(y)} + p_x(y)\frac{\partial \phi_{xy}}{\partial q_y(x)}\right).$$

While Rosen's version requires that diagonal dominance conditions hold globally for all admissible strategies, which is very stringent, the theorem 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 simplexes (for which the index equals 1 and every Karush-Kuhn-Tucker 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 (denoted with a star). Thus, the equilibrium is unique if diagonal dominance conditions hold along the equilibrium path, i.e., if conditions (a) and (b) of the theorem are satisfied. □

The assumption that cost functions are increasing and convex is a natural one. The additional diagonal dominance conditions in our case can be interpreted as implying that the search costs should be sufficiently high for the equilibrium to be unique. If the congestion function has negative derivatives with respect to search intensities, this negativity enlarges the set of parameters under which the equilibrium is unique but in the extreme could lead to non-existence.

If diagonal dominance conditions do not hold in equilibrium, then there can be multiple equilibria. This is a well-known outcome of the frictionless assignment model, which is a special case of our model under search costs approaching zero and no congestion. 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."

The result of Theorem 2 is intuitive. Recall that there are two motives for a female of type x to target a male of type y: the productive and the strategic. 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 on one type of counter-party and exclude all others. It does not matter what portion of the payoff a female of type x will get from a match with a male of type y if the male places a low probability on a female of type 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 type of male and that type of male responds by placing a unit probability on that type of female. Each equilibrium of this kind implements a matching of the classical assignment problem, but as mentioned earlier, not all of them are stable.

As θ increases, 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 matching with a counterparty who 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 frictionless assignment model and the random matching model are special cases of our targeted search model, when θ is either very low or very high.

3.4 Efficiency

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 agents would be infinitely precise.

To study the constrained efficient allocation, we impose on the social planner the same costs of search that are faced by the agents. Let the payoff functions be defined as in (3) and (4). Then, the social planner maximizes the following welfare function:

$$W = \max_{p_x(y)\in S_x, q_y(x)\in S_y} \sum_{y=1}^{M} \mu_y Y_y + \sum_{x=1}^{F} \mu_x Y_x.$$

We can re-write the objective functions of the agents introducing the linear constraints on strategies via Lagrange multipliers (λ_x and λ_y). Then the first-order conditions for optimality are

$$\frac{\partial W}{\partial p_x\left(y\right)} = \Phi_{xy} q_y\left(x\right)\left(\phi_{xy} + p_x\left(y\right)\frac{\partial \phi_{xy}}{\partial p_x\left(y\right)}\right) - \frac{\partial c_x}{\partial \kappa_x}\left(\ln p_x\left(y\right)\sum_{y=1}^{M}\mu_y + 1\right) - \lambda_x = 0, \tag{13}$$

$$\frac{\partial W}{\partial q_y\left(x\right)} = \Phi_{xy} p_x\left(y\right)\left(\phi_{xy} + q_y\left(x\right)\frac{\partial \phi_{xy}}{\partial q_y\left(x\right)}\right) - \frac{\partial c_y}{\partial \kappa_y}\left(\ln q_y\left(x\right)\sum_{x=1}^{F}\mu_x + 1\right) - \lambda_y = 0. \tag{14}$$

As we have established previously, the objective functions of agents are twice continuously differentiable and concave in all the strategies. Hence, first-order conditions of the welfare function with respect to strategies of the agents are necessary and sufficient conditions for an efficient allocation. Rearranging and substituting out Lagrange multipliers, we obtain the following proposition:

Proposition 2. *Under assumptions A.1-A.5, the constrained efficient allocation satisfies*

$$p_x^*\left(y\right) = \frac{exp\left(\frac{\Phi_{xy} q_y^o(x)\left(\phi_{xy}^o + p_x^o(y)\frac{\partial \phi_{xy}^o}{\partial p_x^o(y)}\right)}{\partial c_x/\partial \kappa_x|_{p_x^o(y)}}\right)}{\sum_{y'=1}^{M}\mu_{y'} exp\left(\frac{\Phi_{xy'} q_{y'}^o(x)\left(\phi_{xy'}^o + p_x^o(y')\frac{\partial \phi_{xy}^o}{\partial p_x^o(y')}\right)}{\partial c_x/\partial \kappa_x|_{p_x^*(y')}}\right)}, \tag{15}$$

$$q_y^*\left(x\right) = \frac{exp\left(\frac{\Phi_{xy} p_x^o(y)\left(\phi_{xy}^o + q_y^o(x)\frac{\partial \phi_{xy}^o}{\partial q_y^o(x)}\right)}{\partial c_y/\partial \kappa_y|_{q_y^*(x)}}\right)}{\sum_{x'=1}^{F}\mu_{x'} exp\left(\frac{\Phi_{x'y} p_{x'}^o(y)\left(\phi_{x'y}^o + q_y^o(x')\frac{\partial \phi_{x'y}^o}{\partial q_y^o(x')}\right)}{\partial c_y/\partial \kappa_y|_{q_y^*(x')}}\right)}. \tag{16}$$

The structure of the social planner's solution is very similar to the structure of the decentralized equilibrium. By comparing constrained efficient allocations with the equilibrium allocations from Proposition 1, it becomes clear that competitive equilibria are generally inefficient. From a female's perspective, the difference is that, while the equilibrium strategy depends on her private gain ε_{xy}, the socially optimal strategy depends on the social gain Φ_{xy} adjusted for the effect her strategy has on congestion, $\frac{p_x}{\phi_{xy}}\frac{\partial \phi_{xy}}{\partial p_x}$.

The same difference holds from a male's perspective. Thus, it is socially optimal for both females and males to consider the total payoff adjusted for congestion, while in the decentralized equilibrium they consider only their private payoffs. Inefficiency may arise if an agent increases her search intensity (by choosing a more concentrated strategy), because it generates a positive externality for other agents. Due to the strategic motive, when an agent exerts more effort to better target preferred individuals, it not only increases the agent's probability of being matched but also increases the search effort that other agents have to exert. Agents do not take into account this effect, since their strategy depends solely on the private gain as opposed to the social one.[16]

While the decentralized equilibrium is generally inefficient, there are conditions that can be imposed on the congestion function under which the competitive equilibrium is socially efficient. The following theorem provides such conditions:

Theorem 3. *Under assumptions A.1-A.6, the matching equilibrium is inefficient unless*

1. *males and females are homogeneous or*
2. *the congestion function takes on the form*

$$\phi_{xy}\left(p_x\left(y\right), q_y\left(x\right), \mu_x, \mu_y\right) = \frac{1}{p_x^{\alpha_{xy}}(y) q_y^{1-\alpha_{xy}}(x)} \Psi\left(\mu_x, \mu_y\right)$$

and the parameter α_{xy} equals the surplus split for agents of types x and y:

$$\eta_{xy} = \alpha_{xy}\Phi_{xy}, \qquad \varepsilon_{xy} = \left(1-\alpha_{xy}\right)\Phi_{xy}, \qquad \textit{for all } x, y.$$

The proof is in Appendix A. The theorem states that, with a constant-returns-to-scale technology for the meeting function, if the parameter of the congestion function is equal to the share of the surplus obtained by the partner, the equilibrium is efficient. The aggregate matching function for types x and y becomes a constant-returns-to-scale Cobb-Douglas function of the search intensities, with power parameters equal to the surplus split:

$$\mathfrak{M}_{x,y} = \mu_x \mu_y p_x\left(y\right) q_y\left(x\right) \phi_{xy} = p_x^{\frac{\varepsilon_{xy}}{\Phi_{xy}}} q_y^{\frac{\eta_{xy}}{\Phi_{xy}}} \mu_x \mu_y \Psi\left(\mu_x, \mu_y\right). \tag{17}$$

[16]This result is reminiscent of the holdup problem where there are goods with positive externalities and the producer undersupplies the good if not fully compensated by the marginal social benefits that an additional unit of the good would provide to society. The kind of inefficiency we obtain is similar to that in Mortensen (1982). Mortensen's proposed solution of giving the whole surplus to the partner that initiated the meeting would fail in our environment because both the female and the male initiate the meeting and also because the gains are not easily transferable.

It should be noted that, to create a fully constant-returns-to-scale matching function, we assume $\Psi\left(\mu_x,\mu_y\right)=1$. Indeed, assuming high costs of search, it follows that $p_x\left(y\right)=1/\left(M\mu_y\right)$, $q_y\left(x\right)=1/\left(F\mu_y\right)$; therefore,

$$\Sigma_x\Sigma_y\mathfrak{M}_{x,y}\sim MF\mu_x\mu_y p_x^{\alpha}q_y^{1-\alpha}\sim MF\mu_x\mu_y\left(\frac{1}{M\mu_y}\right)^{\alpha}\left(\frac{1}{F\mu_x}\right)^{1-\alpha}=\left(F\mu_x\right)^{\alpha}\left(M\mu_y\right)^{1-\alpha}.$$

The intuition for the efficiency result is reminiscent of the Hosios (1990) condition, as our model features a positive and a negative externality. The positive externality stems from the fact that an individual with a more targeted strategy incentivizes other people in the market to exert more effort if they want to find a match. This externality is not internalized because the individual surplus differs from the collective surplus. The negative externality stems from the increased congestion in the submarkets (i.e., types) targeted with higher intensity. This externality is not internalized since the meeting rates are given. When the parameter of the congestion function equals the share of the surplus that the agent receives from a match, the two externalities balance out exactly, producing a constrained efficient equilibrium allocation.

An important property of our meeting technology is the assumption of no inter-type congestion externalities, i.e., that strategies of females of type $x'\neq x$ and males of type $y'\neq y$ do not directly affect the meeting rates of females of type x with males of type y. This assumption on the congestion function is important for the social efficiency of competitive equilibrium. If on the contrary the inter-type congestion externalities were present as in Shimer and Smith (2001), then we would, like them, conclude that equilibria are generically inefficient for any functional form of the congestion function.

3.5 Uniqueness under a particular meeting technology

There is at least one particular congestion function that produces a unique equilibrium—regardless of the value of the search costs—that is socially efficient and for which the aggregate matching function features constant returns to scale.

A.7 The congestion function is given by $\phi_{xy}=p_x^{-\alpha_{xy}}q_y^{-\beta_{xy}}$.

The following theorem places additional restrictions on the congestion function and surplus split to achieve uniqueness in the decentralized economy:

Theorem 4. *Under assumptions A.1-A.7, the matching equilibrium is unique if*

a. *the congestion function is the same for all pairs of types:*

   for all $x \in \{1, ..., F\}, y \in \{1, ..., M\}$, *we have* $\alpha_{xy} = \alpha$, $\beta_{xy} = \beta$;

b. *the surplus split is the same for all pairs of types:*

   for all $x, x' \in \{1, ..., F\}$, $y, y' \in \{1, ..., M\}$, *we have*

   $\frac{\varepsilon_{xy}}{\Phi_{xy}} = \frac{\varepsilon_{x'y}}{\Phi_{x'y}} = \frac{\varepsilon_{xy'}}{\Phi_{xy'}}$, $\frac{\eta_{xy}}{\Phi_{xy}} = \frac{\eta_{x'y}}{\Phi_{x'y}} = \frac{\eta_{xy'}}{\Phi_{xy'}}$; *and*

c. *the coefficients of the congestion function satisfy*

   $\alpha > 0, \beta > 0$, $\alpha + \beta \geq 1$, $\min(\alpha, \beta) < 1$.

The proof is in Appendix A. Note that the conditions that deliver uniqueness of the decentralized equilibrium do not impose any additional restrictions on the costs of search so long as they are positive, increasing, and convex.

To conclude, Theorem 4 is the final step in establishing that, if the congestion function has a special functional form $\phi_{xy}(p_x, q_y, \mu_x, \mu_y) = (p_x)^{-\alpha}(q_y)^{-(1-\alpha)}$, the surplus is split proportionally as $\frac{\varepsilon_{xy}}{\Phi_{xy}} = 1-\alpha$, and the parameter α is the same for all pairs of types (x, y), then the competitive equilibrium exists, is unique, and is constrained efficient for any positive search costs. Moreover, the aggregate matching function exhibits constant returns to scale. These properties will prove extremely useful for empirical analysis.

3.6 Sorting: 2×2 case

In this subsection we study the sorting properties of the equilibria that the model yields in the 2×2 case. Consider the case where there are $F = 2$ types of females (high and low), $M = 2$ types of males (high and low), and μ individuals of each type. Consider a symmetric payoff function $\Phi_{xy} = 2\begin{bmatrix} u & 1 \\ 1 & d \end{bmatrix}$, where the first row represents the high-type female, the second row the low-type female, the left column the high-type male, and the right column the low-type male. This structure implies that matches between high-type women and low-type men get the same total payoff as matches between low-type women and high-type men.

Following results from the previous section, assume that upon matching, the surplus is split evenly between males and females and the special congestion function

$\phi(p,q,\mu) = (pq)^{-1/2}$ guarantees both uniqueness and efficiency of equilibria. Assume that search costs are linear in effort and that the marginal costs are the same for all agents, $\theta_x = \theta_y = \theta$.

Under these assumptions, the problems of the high-type male and high-type female are identical, as are the problems of the low types. Denote the probability that a high-type female places on a high-type male by $p_h(h)$, and denote the probability that a low-type female places on a high-type male by $p_l(h)$. These two choice variables determine all the remaining equilibrium objects, due to symmetry and the requirement that probability distributions sum up to 1. Using the necessary conditions for equilibria from Proposition 1, we can simplify further to obtain the following two equations:

$$\mu p_h(h)\left(1+\exp\left(\frac{1}{\theta}\left(\sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}}-u\right)\right)\right)=1, \tag{18}$$

$$\mu p_l(h)\left(1+\exp\left(\frac{1}{\theta}\left(d-\sqrt{\frac{1-\mu p_h(h)}{\mu p_l(h)}}\right)\right)\right)=1. \tag{19}$$

These two equations represent the "best responses" of high-type and low-type agents to each others' strategies. We illustrate the shapes of these strategies in Figure 2.[17]

When in equilibrium high types choose $p_h(h) > 1/(2\mu)$, i.e., target high types, and low types choose $p_l(h) < 1/(2\mu)$, i.e., target low types, we call this positive assortative matching (PAM). All PAM equilibria are located in Quadrant IV of Figure 2. Similarly, when high types choose to target low types and low types choose to target high types, we call this negative assortative matching (NAM). All NAM equilibria are located in Quadrant II of Figure 2. The remaining equilibria, where everybody targets high types (Quadrant I) or everybody targets low types (Quadrant III), we call *mixing equilibria.*

In what follows we establish conditions under which each type of equilibria prevails, in two extreme cases, when marginal costs of search approach zero and when they approach infinity. In summary, we show that if preferences are vertical—i.e., everyone gets a higher payoff if matched with a high type—the equilibrium can be PAM, mixing, or NAM depending on the cardinal properties of the surplus in the zero-cost limit, and

[17] For parameters $\mu = 1$, $\theta = e^{-0.4}$, $d = 0.83$, $u = 1.34$. An interactive version of this graph may be accessed by the reader at https://www.desmos.com/calculator/yvplpyoesg.

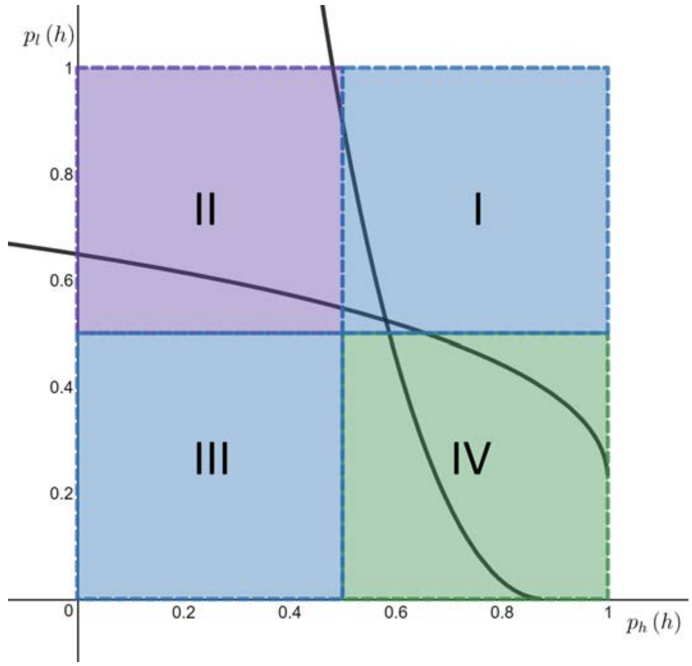

Figure 2: Best Responses and Types of Equilibria.

the equilibrium is mixing in the high-cost limit. If preferences are horizontal—i.e., each type gets a higher payoff if matched with someone of their same type—the equilibrium is PAM for all positive finite values of costs.

Proposition 3. *Sorting properties:*

1. *If $u > 1 > d$. As $\theta \to 0$:*
   a. *PAM: if $u > \frac{1}{d}$, then $\mu p_h(h) \to 1$, $\mu p_l(h) \to 0$;*
   b. *Mixing: if $\sqrt{2} \leq u < \frac{1}{d}$, then $\mu p_h(h) \to 1 - u^{-2}$, $\mu p_l(h) \to 1$;*
   c. *NAM: if $u < \frac{1}{d}$, $u < \sqrt{2}$, then $\mu p_h(h) \to 1 - u^{-2}$, $\mu p_l(h) \to 1$;*
2. *PAM: If $u > 1$, $d > 1$. As $\theta \to 0$, $\mu p_h(h) \to 1$, $\mu p_l(h) \to 0$.*
3. *As $\theta \to \infty$, $\mu p_h(h) \to \frac{1}{2}$, $\mu p_l(h) \to \frac{1}{2}$ (random matching):*
   a. *Mixing: If $u > 1 > d$, then $\mu p_h(h) > \frac{1}{2}$, $\mu p_l(h) > \frac{1}{2}$;*
   b. *PAM: If $u > 1, d > 1$, then $\mu p_h(h) > \frac{1}{2}$, $\mu p_l(h) < \frac{1}{2}$.*

The proof is in Appendix A. A common result in the search and matching literature is that matching is positively assortative if the surplus function has supermodular properties. Note that in the 2×2 example the surplus is supermodular if $u + d > 2$ and log-supermodular if $ud > 1$. The first condition seems unrelated to our sorting results, but the second indeed determines the sorting pattern. In fact, it follows from Proposition 3 that, in this specific 2×2 case, whenever the surplus is log-supermodular, we get a PAM equilibrium for $\theta \to 0$. However, we can also get a mixing equilibrium for high-enough costs when the payoffs are log-supermodular (see Proposition 3.3a).

In the zero-cost limit we also obtain NAM if $ud < 1$, $u < \sqrt{2}$, and we get the new type of mixing equilibrium for intermediate values of parameters in between PAM and NAM equilibria. This suggests that when approaching the frictionless limit, the sorting patterns might be similar in spirit to those obtained under frictionless assignment (Becker (1973)) or in frictional models (Shimer and Smith (2000), Smith (2006)). Although, we have not been able to prove in general that equilibria exhibit PAM in the zero-cost limit when preferences are log-supermodular, we have not been able to find a counter-example to this statement. Nevertheless, we find that for non-limit cases, log-supermodularity is not a useful guide for predicting sorting in equilibrium. Instead, mixing-type equilibria become increasingly common with the increase in the number of types and for non-zero search costs. We discuss how we evaluate the types of equilibria and the shapes of surplus in the next section.

There also seems to be no direct link between log-supermodularity and the horizontality or verticality of preferences. In the 2×2 case, preferences are horizontal when $u > 1$, $d > 1$ and vertical when $u > 1$, $d < 1$. In the zero-cost limit, as long as $ud > 1$, we obtain PAM regardless of whether preferences are horizontal ($d > 1$) or vertical ($1 > d > \frac{1}{u}$).

4 Empirical Application

In this section we explain how the model can be used to recover underlying preferences; we also illustrate how, due to the presence of the strategic motive for which relative importance is determined endogenously, the degree of horizontality of the recovered payoff function can be different from that estimated using models where there is no strategic component or where the relative importance of this motive is built-in ad hoc. We use these results to check for examples of the mixing equilibrium found in our theoretical framework.

To this purpose, in Subsection 4.1 we describe our identification strategy; in Subsection 4.2 we define a formal way of measuring both the degree of horizontality of preferences and the degree of assortativeness of matching; and in Subsection 4.3 we show the results of the calibration exercise, applying results from Subsection 4.2. In Subsection 4.4 we present a stark example for which our model is particularly empirically relevant. Bruch and Newman (2018) find that when searching for a partner, individuals pursue partners who are above their league in some characteristics. We show that this is a behavior that derives naturally from our environment.

4.1 Identification

To calibrate our model, we need data on contact rates or on the number of matches. In what follows, we describe how to recover underlying preferences using matching data (see Appendix C for the methodology using contact rates).[18]

One can observe matching rates—$\mathcal{M}_{x,y}$— the numbers of men and women married to each other in a given period of time, as well as the number of men and women μ_x and μ_y that remained single over that period. In this case, the data contain observations with $M \times F$ degrees of freedom. Of course, there are not enough restrictions in the data to identify payoffs in a non-transferable utility case. But assuming transferable utility with a predefined split of the joint payoff between males and females, Φ_{xy}, the unobserved payoff functions also have $M \times F$ independent parameters. In this case, it is possible to identify the payoff function from just matching rates. Taking this into consideration, for our empirical methodology, we make the following identifying assumptions:

I.1 Search costs are identical across females and males, $\theta_x = \theta_y = \theta$.

[18] The contact rate data used in Hitsch et al. (2010) and in Bruch and Newman (2018) is confidential, and we have not been able to get access to it.

I.2 Search costs are normalized to $\theta = 1$.

I.3 Payoff Φ_{xy} is split equally between males and females; i.e., $\epsilon_{xy} = \frac{1}{2}\Phi_{xy}$.

I.4 The congestion function is given by $\phi_{xy} = 1/\sqrt{p_x(y)\, q_y(x)}$.

Assumptions I.1-I.2 allow us to focus on identifying the ratio Φ_{xy}/θ as the object entering equilibrium conditions since the payoff and marginal cost cannot be identified separately. Assumptions I.3-I.4 guarantee existence, uniqueness, and efficiency of equilibrium, as we have shown in Section 3.[19] Under these assumptions we do not need to worry about multiplicity of equilibria. Equilibrium uniqueness implies a one-to-one mapping from the shape of the payoff function Φ_{xy} to matching rates $\mathfrak{M}_{x,y}$ conditional on the number of agents μ_x, μ_y.

Building on these identifying assumptions, we write down a loss function and minimize it to obtain the surplus. The loss function is defined as follows:

$$\mathscr{L}\left(\mathcal{M}_{x,y}|\Phi_{xy}, \mu_x, \mu_y\right) = -\sum_{x\in\{1,\dots,F\}}\sum_{y\in\{1,\dots,M\}} \mathcal{M}_{x,y}\left(\ln \mathfrak{M}^*_{x,y}\left(\Phi_{xy}, \mu_x, \mu_y\right) - \ln \mathcal{M}_{x,y}\right),$$

where $\mathfrak{M}^*_{x,y}\left(\Phi_{xy}\right)$ is the equilibrium matching rate produced by the model for surplus Φ_{xy}, according to equation (17).

The loss function is minimized when the model matches the data perfectly; i.e., $\mathfrak{M}^*_{x,y}\left(\Phi_{xy}, \mu_x, \mu_y\right) = \mathcal{M}_{x,y}$, and the minimized loss function yields $\hat{\Phi}_{xy} = \arg\min_{\Phi_{xy}} \mathscr{L}$. In order to find the global minimum in a high-dimensional space, we initialize the algorithm at a large number of different initial points and then compare the points to which the search algorithm converged. See Appendix B for a derivation of a closed-form solution and a discussion of identification in the 2×2 case and more generally.

The identifying assumptions that we make, by guaranteeing uniqueness of equilibria, make the calibration procedure very well-behaved. We find that under our identifying assumptions the computational difficulty in recovering the parameters of the model is the same as that of solving for the matching equilibrium.

[19] Note that I.4 incorporates the assumption of no direct inter-type congestion externalities, i.e., that the meeting rate between women of type x and men of type y is unaffected by the strategies of other types. Assuming external effects generically leads to non-uniqueness with different equilibria exhibiting different sorting patterns, thus making it hard to identify preferences.

4.2 Vertical versus horizontal preferences

In the 2×2 case of the model, it is straightforward to identify its sorting properties and whether preferences are vertical or horizontal. However, for a large number of types of males and females, we need to define how we measure the verticality/horizontality of preferences and how we measure the degree of sorting that characterizes the competitive equilibrium of the model.

We define preferences to be vertical if every type's best match is the same type, and we define preferences to be horizontal if every type's best match is a different type. For preferences that are neither strictly vertical nor strictly horizontal, we can define a "horizontality index" (HI). Let $\omega_x = \left| \{\arg\max_y (\varepsilon_{xy})\}_{x \in \{1,...,F\}} \right| \in \{1, ..., M\}$ be the number of different types of males who are best matches for at least one type of female. Similarly, let $\omega_y = \left| \{\arg\max_x (\eta_{xy})\}_{y \in \{1,...,M\}} \right| \in \{1, ..., F\}$ be the number of different types of females who are best matches for at least one type of male. Then the HI is just a normalized average of the two numbers:

$$H(\varepsilon_{xy}, \eta_{xy}) = (\omega_x + \omega_y - 2) / (M + F - 2).$$

When preferences are vertical, the HI takes a value of zero. When preferences are horizontal, the HI takes a value of 1.

Likewise, we can define an "assortativeness index" (AI) to characterize sorting in equilibrium. Let us denote by $\mathbb{P}_{xy} = [p_x(y)]$ the matrix of all female strategies and by $\mathbb{Q}_{yx} = [q_y(x)]$ the matrix of male strategies. Then let $\xi_x = \left| \{\arg\max_y (\mathbb{P}_{xy})\}_{x \in \{1,...,F\}} \right| \in \{1, ..., M\}$ be the number of different types of males that females target, and let $\xi_y = \left| \{\arg\max_x (\mathbb{Q}_{yx})\}_{y \in \{1,...,M\}} \right| \in \{1, ..., F\}$ be the number of different types of females that males target. Then the AI of an equilibrium is

$$A(\mathbb{P}_{xy}, \mathbb{Q}_{yx}) = (\xi_x + \xi_y - 2) / (M + F - 2).$$

For an assortative equilibrium, the AI equals 1, while for a mixing equilibrium the AI equals zero.

The conceptual difference between the HI and the AI is as follows. The HI characterizes ex-ante preferences of individuals (the shape of the underlying payoff function), that is, whom they target based on the productive motive and abstracting from the strategic one. The AI characterizes the ex-post realization of who matches whom, that is, whom males and females target when taking into account both the productive and the strategic motives.

Note that when preferences are horizontal, the productive and the strategic motives pull in the same direction, meaning that the productive motive will drive individuals to target those with whom the payoff is higher and the strategic motive will drive individuals to target those same people because there is no competition to match with them, given that everyone is better off with a different type. However, when preferences are vertical, the productive motive drives everyone to target the same type—with whom they would get a higher payoff—but the strategic motive will pull them in a different direction to maximize the chance of reciprocation.

Because of this dynamic, when preferences are vertical, the strategic motive makes it such that by looking at the shape of the observed matching rate one cannot directly infer the shape of the underlying payoff function. In a model of unobserved characteristics—like Choo and Siow (2006)—by construction, the shape of the matching rate is the same as the shape of the underlying preferences (there is no strategic motive). To make this point, in the empirical results subsection, we report the AI and HI generated by our model, as well as the HI calculated using the underlying preferences estimated using the Choo and Siow model (HICS). This will illustrate how different the estimated underlying preferences can be depending on whether the model incorporates a strategic motive.

4.3 Empirical results

To calibrate our model, we use data on matching rates in the U.S. for the year 2012 from the Integrated Public Use Microdata Series (IPUMS).[20] We take unmarried males and females and (newly) married couples and assign both males and females to bins corresponding to types in the model. We consider four dimensions along which males and females evaluate each other in the marriage market: income, age, education, and

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

race.

An advantage of our methodology is that it allows for multidimensional characteristics. This is important because younger people are likely to have lower levels of education and lower levels of income. Therefore, one could use more-easily observable characteristics, such as age, to screen for less-easily observable characteristics, such as education and income. If that is indeed what is going on, we might confuse preferences for age with preferences for income and education. To identify the underlying preferences controlling for correlations between characteristics, one needs to be able to estimate the model using multidimensional types.[21]

In each case we choose the cutoffs between bins in such a way as to split the whole sample, representative of the U.S. population, into equally sized bins to the extent possible. We restrict our attention to adults between the ages of 21 and 40. 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. We call newly married the couples that got married in the past 12 months.

We allow three bins for age (22, 27, and 34), three bins for income (low, medium, and high), and two bins for education (school or college). Altogether there are 18 different combinations of these attributes. We calibrate this multidimensional model using the representative sample of the U.S. population for all married couples in the past year and unmarried males and females ages 21-40 in the year 2012.

Figure 3 displays the results. There are four panels: the first two panels represent the strategies of males and females,[22] respectively, and the second two the number of matches and the payoff function, respectively. Note that the number of matches is observed from the data and the strategies and payoff are recovered from the model using the identifying assumptions from Subsection 4.1.

[21]We present the unidimensional results for income, age, education, and race in Appendix D.

[22]Each bar represents $\mu_x p_x(y)$ for the females and $\mu_y q_y(x)$ for the males and, hence, $\sum_y \mu_x p_x(y) = 1$, $\sum_x \mu_y q_y(x) = 1$.

From the payoff function we can see that females are indifferent between 27- and 34-year-old males, but have a strong preference for high-income college-educated males. Male preferences are more heterogeneous; however, they tend to target 27- or 34-year-old females with a college degree and high income. Females have the same targeting strategies as males but with less diversification.

This is an example of a mixing equilibrium as the assortativeness index is 0.32, and preferences are vertical as the horizontality index is 0.11. Overall, desirability increases with income and education, and partners over 25 years of age are preferred by both males and females. This finding is consistent with our unidimensional findings (see Appendix D) for income and education but a little different for age. This finding also suggests that age may be serving as a signal of income and education, rather than agents exhibiting strong preferences over age alone.

If we recover preferences using the Choo and Siow model, we would find the horizontality index to be 0.41, which would point toward preferences that are much more horizontal than those produced by our model. For all of our calibrations—unidimensional or multidimensional—we conclude that preferences are close to vertical.[23]

4.4 People pursue partners "out of their league"

In a study about romantic courtship in online dating markets, Bruch and Newman (2018) find that both males and females, on average, pursue partners who are 25 percent above their own rank. In the author's language, both males and females seem to be "reaching up the desirability ladder," i.e., "pursuing partners who are on average more desirable than themselves," and sometimes even "out of their league."

In this subsection, we show that this behavior derives naturally from our environment.[24] We adopt from Bruch and Newman (2018) the concept of desirability rank.

[23]This finding raised some concern for us in terms of whether this conclusion could be the result of model misspecification. To check whether model misspecification could bias the horizontality of preferences, we conducted a set of Monte-Carlo exercises. We find that even substantial misspecification of the parameters governing the congestion function and the split of the surplus does not seem to bias inference regarding the horizontality or verticality of preferences.

[24]Bruch and Newman reach this conclusion using disaggregated data from a dating site that they don't disclose. As was mentioned earlier, we only have marriage data from IPUMS. However, it is

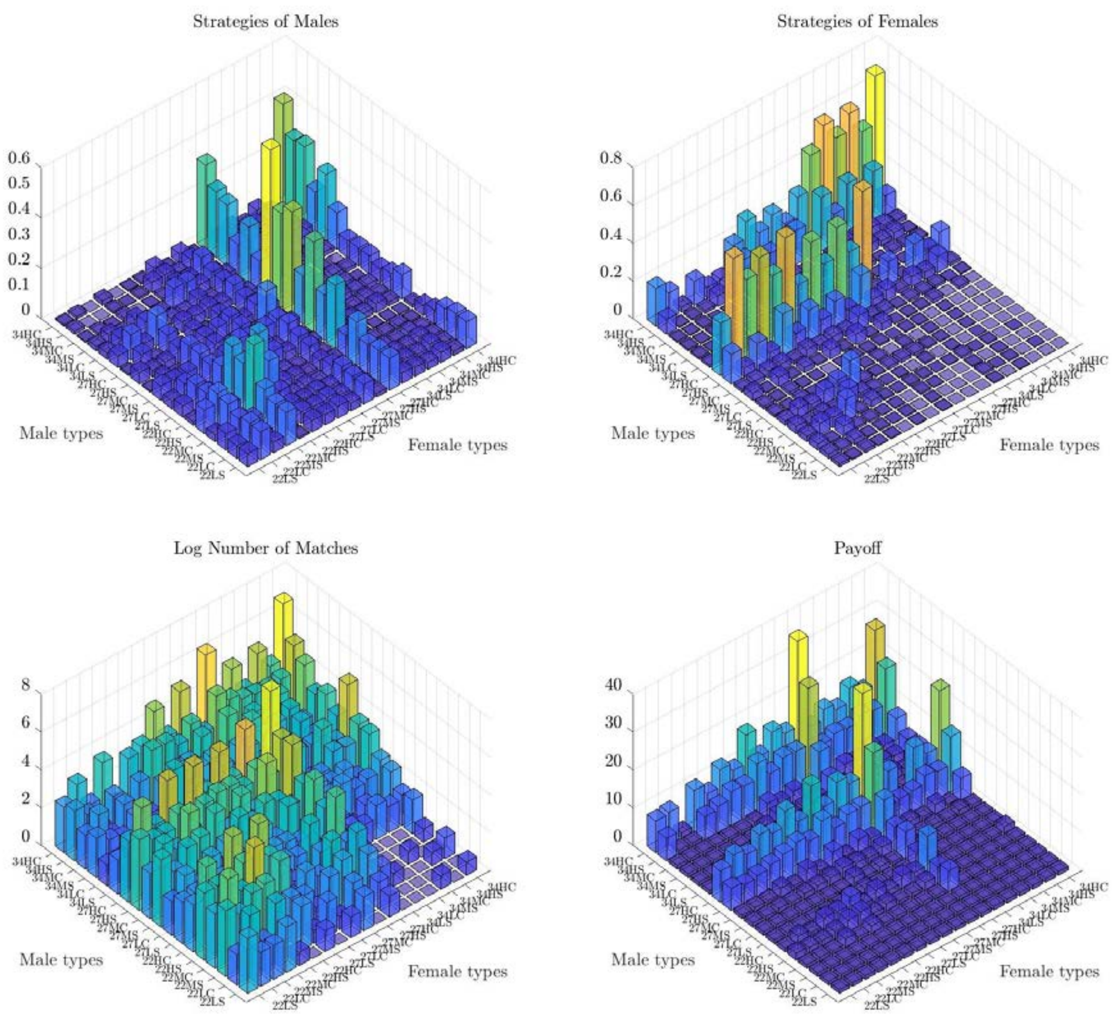

Figure 3: Sorting by Multidimensional Characteristics

Making an analogy between our model and a dating website—used in Bruch and Newman—the estimated competitive equilibrium of our model predicts the relative frequencies at which different types of females would send messages to different types of males and vice versa. We assume that each agent in the model sends an equal and fixed number of messages to agents of the opposite sex. Following the methodology of Bruch and Newman (2018), we use Google's PageRank algorithm to compute the desirability rank for each agent in our model. The algorithm ranks agents based on the number of messages they received, weighted by the rank of the sender. We then compute the average desirability rank for agents of each type. This methodology allows us to compute a desirability index (DI) for each type and see how desirable each type's targets are relative to their own.

In Appendix D we show the DIs as well as the predominant target behavior of each type when sorting by income, age, education, and race. Figures 8 to 12 in Appendix D show the targets as well as the DIs for the results that we report here.

Based on those results, we want to see whether—in our model—individuals target people above or below their own DI (above or below their "league"). To this end we compute the following league indices (LIs) for females and males: For females,

$$LI_x = \sum_{y=1}^{M} \mu_y p_x(y) DI_y - DI_x,$$

where DI_x is the DI of a female of type x. For males,

$$LI_y = \sum_{x=1}^{F} \mu_x q_y(x) DI_x - DI_y,$$

where DI_y is the DI of a male of type y.

A value of zero means that individuals target, on average, people in their own league, a negative value means that they target people below their league, and a positive value means that they target people above their league. Table 1 shows averages over types when sorting by different characteristics. When sorting by education, on average, males

natural to believe that if individuals have this behavior when looking for someone to date, they will have a similar behavior when looking for someone to marry.

	Males	Females	Average
Education	0.27	0.05	0.17
Race	0.19	-0.03	0.09
Income	0.21	0.14	0.18
Age	0.21	0.10	0.16
Inc x Age x Educ	0.30	0.26	0.28

Table 1: Individuals target someone above their “league” (league index)

target females who are 27% above their league, while females target males who are only 5% above. When sorting by race, on average, males target females who are 19% above their league, while females target males who are 3% below their league. When sorting by income, males target females who are 21% above their league while females target males who are 14% above. When sorting by age, on average, males overshoot by 21% and females by 10%.

When we look at the results sorting by multidimensional characteristics, we see that our model predicts that people pursue partners who are, on average, 28% above their own “league.”

Our results are very close to those in Bruch and Newman (2018), who find that individuals pursue others who are on average 25% above their league. In our model this arises because, since agents know that they and others cannot locate their preferred type with accuracy, there is some positive probability that someone above their league will reciprocate their interest.

5 Final Remarks

In this paper we propose a model of probabilistic choice by agents where the degree of randomness in matching is endogenous because deciding whom to contact and locating the best partners involves search effort. Agents' optimal targeting strategy resulting from the model balances a productive motive whereby agents contact potential partners who will render a higher payoff and a strategic motive that drives agents toward potential partners who are more likely to reciprocate their interest. We find that accounting for the interaction between strategic and productive considerations aids identification of underlying preferences, while ignoring this interaction may result in misleading implications regarding the degree of mismatch and hence the welfare losses associated with it. Understanding who meets whom is crucial for understanding who marries whom and who should marry whom instead.

The model we present in this paper is deliberately static for simplicity and clarity of exposition. Nevertheless, we believe that complications arising in a repeated setting due to additional frictions associated with time costs of search are empirically relevant. We describe an extension of our model to a repeated setting with time costs of search in Appendix E and discuss its empirical relevance. Another important extension that we do not touch is the dynamic change in the number of searching agents due to match formation and due to search while married. We leave the study of a fully dynamic repeated setting for future research.

We develop an empirical methodology and apply it to aggregate data on the U.S. marriage market. Aggregate data allow identification of preferences only under stringent assumptions on the division of surplus between men and women. To relax these assumptions it would be useful to estimate our model using data on contact rates, such as in Bruch and Newman (2018) or Hitsch et al. (2010). Another interesting line of enquiry would be to combine data on matching rates and contact rates to study the properties of the meeting technology.

Even though our application focuses on the U.S. marriage market, our model is well-suited to study a host of real-life matching markets where the potential matches are numerous and diverse. Matching in labor, education, and health care are just a few examples of markets in which agents need to exert considerable effort to screen long lists of heterogenous candidates. Moreover, all these markets display match formation between superior and inferior types that can be rationalized by our model and evaluated with our methodology.

Finally, 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 nontransferable utility." *Journal of Economic Theory*, 113, pp. 182-198.

[2] **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.

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

[4] **Bruch, E. E. and M. E. J. Newman (2018).** "Aspirational Pursuit of Mates in Online Dating Markets" *Science Advances*, 4(8), pp. 1-6.

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

[6] **Chade, H. (2001).** "Two-Sided Search and Perfect Segregation with Fixed Search Costs." *Mathematical Social Sciences*, 42, 31-51.

[7] **Chade, H., Eeckhout, J. and L. Smith (2017).** "Sorting Through Search and Matching Models in Economics." *Journal of Economic Literature*, 55(2), 493-544.

[8] **Cheremukhin, A., Popova A. and A. Tutino (2015).** "A Theory of Discrete Choice with Information Costs." *Journal of Economic Behavior and Organization*, 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] **Gale, D. and H. Nikaido (1965).** "The Jacobian Matrix and Global Univalence of Mapping." *Mathematische Annalen*, 159, pp. 81-93.

[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 (2019).** "Optimal Taxation, Marriage, Home Production, and Family Labor Supply." *Econometrica*, 87(1), pp. 291-326.

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

[17] **Hosios, A. J. (1990).** "On the Efficiency of Matching and Related Models of Search and Unemployment." *Review of Economic Studies*, 57(2), pp. 279-298.

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

[19] **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.

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

[21] **Mortensen, D. T. (1982).** "Property Rights and Efficiency in Mating, Racing, and Related Games," *American Economic Review*, 72(5), pp. 968-979.

[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. and L. Smith (2001).** "Matching, Search and Heterogeneity." *The B.E. Journal of Macroeconomics*, 1(1), pp. 1-18.

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

[26] **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.

[27] **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.

[28] **Smith, L. (2006).** "The Marriage Model with Search Frictions," *Journal of Political Economy*, 114(6), pp. 1124-1146.

[29] **Tarski, A. (1955).** "A Lattice-theoretical Fixpoint Theorem and Its Applications," *Pacific Journal of Mathematics*, 5, pp. 285-309.

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

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

Appendix A: Omitted Proofs

A.1 Proof of Theorem 3

Proof. The proof proceeds in three steps:

Step 1

As we have already discussed, 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 CE the first-order conditions of the decentralized equilibrium and by PO the first-order conditions of the social planner. In formulae,

$\mathrm{PO}_{p_x(y)}$: $\Phi_{xy}\tilde{q}_y(x)\left(\tilde{\phi}_{xy}+\tilde{p}_x(y)\frac{\partial\tilde{\phi}_{xy}}{\partial\tilde{p}_x(y)}\right)-\frac{\partial c_x}{\partial\tilde{\kappa}_x}\left(\ln\tilde{p}_x(y)\sum_{y=1}^{M}\mu_y+1\right)=\tilde{\lambda}_x$

$\mathrm{PO}_{q_y(x)}$: $\Phi_{xy}\tilde{p}_x(y)\left(\tilde{\phi}_{xy}+\tilde{q}_y(x)\frac{\partial\tilde{\phi}_{xy}}{\partial\tilde{q}_y(x)}\right)-\frac{\partial c_y}{\partial\tilde{\kappa}_y}\left(\ln\tilde{q}_y(x)\sum_{x=1}^{F}\mu_x+1\right)=\tilde{\lambda}_y$

$\mathrm{CE}_{p_x(y)}$: $\varepsilon_{xy}q_y(x)\phi_{xy}-\frac{\partial c_x}{\partial\kappa_x}\left(\ln p_x(y)\sum_{y=1}^{M}\mu_y+1\right)=\lambda_x$

$\mathrm{CE}_{q_y(x)}$: $\eta_{xy}p_x(y)\phi_{xy}-\frac{\partial c_y}{\partial\kappa_y}\left(\ln q_y(x)\sum_{x=1}^{F}\mu_x+1\right)=\lambda_y$.

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.

Imagine that the two conditions above hold. Then, by construction,

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

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

where we denote marginal costs by a_y and a_x for males and females, respectively. It then follows that

$$\Phi_{xy}\tilde{p}_x(y)\left(\tilde{\phi}_{xy}+\tilde{q}_y(x)\frac{\partial\tilde{\phi}_{xy}}{\partial\tilde{q}_y(x)}\right)-\tilde{\lambda}_y=\eta_{xy}p_x(y)\phi_{xy}-\lambda_y,$$

$$\Phi_{xy}\tilde{q}_y(x)\left(\tilde{\phi}_{xy}+\tilde{p}_x(y)\frac{\partial\tilde{\phi}_{xy}}{\partial\tilde{p}_x(y)}\right)-\tilde{\lambda}_x=\varepsilon_{xy}q_y(x)\phi_{xy}-\lambda_x$$

for all x and y. We can use the first-order conditions of males to derive the formulae for λ_y and $\tilde{\lambda_y}$ and substitute them back in. We conclude that

$$\Phi_{xy}p_x(y)\left(\phi_{xy}+q_y(x)\tfrac{\partial\phi_{xy}}{\partial q_y(x)}\right)-\eta_{xy}p_x(y)\phi_{xy}=$$
$$=\Phi_{x'y}p_{x'}(y)\left(\phi_{x'y}+q_y(x')\tfrac{\partial\phi_{x'y}}{\partial q_y(x')}\right)-\eta_{x'y}p_{x'}(y)\phi_{x'y}\qquad\text{for all } x \text{ and } x'$$
$$\Phi_{xy}q_y(x)\left(\phi_{xy}+p_x(y)\tfrac{\partial\phi_{xy}}{\partial p_x(y)}\right)-\varepsilon_{xy}q_y(x)\phi_{xy}=$$
$$=\Phi_{xy'}q_{y'}(x)\left(\phi_{xy'}+p_x(y')\tfrac{\partial\phi_{xy'}}{\partial p_x(y')}\right)-\varepsilon_{xy'}q_{y'}(x)\phi_{xy'}\qquad\text{for all } y \text{ and } y'.$$

Therefore, either

a. $\Phi_{x'y'}=\Phi_{x''y'}=\Phi_{x'y''}$, $\eta_{x'y'}=\eta_{x''y'}=\eta_{x'y''}$ and $\varepsilon_{x'y'}=\varepsilon_{x''y'}=\varepsilon_{x'y''}$ for all x',x'', y' and y'', or

b. $\Phi_{xy}\left(\phi_{xy}+q_y(x)\frac{\partial\phi_{xy}}{\partial q_y(x)}\right)=\eta_{xy}\phi_{xy}$ and $\Phi_{xy}\left(\phi_{xy}+p_x(y)\frac{\partial\phi_{xy}}{\partial p_x(y)}\right)=\varepsilon_{xy}\phi_{xy}$ for all x and y.

Case a) implies that there is no heterogeneity among types, and strategies of all agents are uniform, which indeed constitutes a socially optimal random matching equilibrium.

Case b) would require that

$$\frac{q_y(x)}{\phi_{xy}}\frac{\partial\phi_{xy}}{\partial q_y(x)}=\frac{\eta_{xy}}{\Phi_{xy}}-1=-\frac{\varepsilon_{xy}}{\Phi_{xy}},$$
$$\frac{p_x(y)}{\phi_{xy}}\frac{\partial\phi_{xy}}{\partial p_x(y)}=\frac{\varepsilon_{xy}}{\Phi_{xy}}-1=-\frac{\eta_{xy}}{\Phi_{xy}}.$$

This system of differential equations with respect to $\phi(p,q)$ has a unique solution of the form

$$\phi(p_x,q_y,\mu_x,\mu_y)=p_x^{-\frac{\eta_{xy}}{\Phi_{xy}}}q_y^{-\frac{\varepsilon_{xy}}{\Phi_{xy}}}\Psi(\mu_x,\mu_y).$$

$\square$

A.2 Proof of Theorem 4

Proof. To establish uniqueness, we use the univalence theorem of Gale and Nikaido (1965), which requires the determinant of the Hessian matrix to be negative semi-definite (i.e., the matrix $H + H^T$ to be negative definite) for all p_x and q_y. Exploiting the block structure of the Hessian matrix as in Theorem 1, the sufficient condition required by Gale and Nikaido's theorem is satisfied if

$$4\left|\frac{\partial^2 Y_x}{\partial p_x \partial p_x}\right|\left|\frac{\partial^2 Y_y}{\partial q_y \partial q_y}\right| > \left(\frac{\partial^2 Y_x}{\partial p_x \partial q_y} + \frac{\partial^2 Y_y}{\partial q_y \partial p_x}\right)^2$$

for all $p_x(y) \in S_x$, $q_y(x) \in S_y$, for all $y \in \{1, ..., M\}$ and for all $x \in \{1, ..., F\}$.

Let $a_x > 0$ and $a_y > 0$ be some scaling factors associated with the payoffs Y_x and Y_y, respectively, for $y \in \{1, ..., M\}$ and $x \in \{1, ..., F\}$. Taking the derivatives of the first-order conditions with the scaling factors and substituting into the above expression, it follows that

$$4a_x a_y \left(\Gamma_x - \varepsilon_{xy} q_y(x) \frac{\partial \phi_{xy}}{\partial p_x(y)}\right)\left(\Gamma_y - \eta_{xy} p_x(y) \frac{\partial \phi_{xy}}{\partial q_y(x)}\right) >$$

$$> \left(a_x \varepsilon_{xy}\left(\phi_{xy} + q_y(x)\frac{\partial \phi_{xy}}{\partial q_y(x)}\right) + a_y \eta_{xy}\left(\phi_{xy} + p_x(y)\frac{\partial \phi_{xy}}{\partial p_x(y)}\right)\right)^2,$$

where

$$\Gamma_x = \frac{\partial c_x}{\partial \kappa_x}\frac{1}{p_x(y)} + \frac{\partial^2 c_x}{\partial \kappa_x \partial \kappa_x}\left(\ln p_x(y) \sum_{y=1}^{M} \mu_y + 1\right)^2$$

and

$$\Gamma_y = \frac{\partial c_y}{\partial \kappa_y}\frac{1}{q_y(x)} + \frac{\partial^2 c_y}{\partial \kappa_y \partial \kappa_y}\left(\ln q_y(x) \sum_{x=1}^{F} \mu_x + 1\right)^2.$$

Leaving the strictly positive terms that involve search costs on the left-hand side and moving the remaining term to the right-hand side, it follows that the Hessian is negative semi-definite if there exists a pair of weights (a_x, a_y) such that the following quadratic expression equals zero for any admissible strategies $p_x(y)$ and $q_y(x)$:

$$\left(a_x\left(\phi_{x,y}+q_y\frac{\partial\phi_{xy}}{\partial q_y}\right)\varepsilon_{xy}+a_y\left(\phi_{x,y}+p_x\frac{\partial\phi_{xy}}{\partial p_x}\right)\eta_{xy}\right)^2-4a_xa_y\frac{\partial\phi_{xy}}{\partial p_x}\frac{\partial\phi_{xy}}{\partial q_y}p_xq_y\varepsilon_{xy}\eta_{xy}=0.$$

Recall that under assumption A.7, $p_x\frac{\partial\phi_{xy}}{\partial p_x}=-\alpha_{xy}\phi_{xy}$ and $q_y\frac{\partial\phi_{xy}}{\partial q_y}=-\beta_{xy}\phi_{xy}$. Therefore, the quadratic form simplifies to

$$\left(\left(1-\beta_{xy}\right)\left(a_x\phi_{xy}\varepsilon_{xy}\right)+\left(1-\alpha_{xy}\right)\left(a_y\phi_{xy}\eta_{xy}\right)\right)^2-4\alpha_{xy}\beta_{xy}\left(a_x\phi_{xy}\varepsilon_{xy}\right)\left(a_y\phi_{xy}\eta_{xy}\right)=0.$$

We can then define a new variable $t_{xy}=\sqrt{\frac{a_x\varepsilon_{xy}}{a_y\eta_{xy}}}>0$ that simplifies the quadratic form as follows:

$$\left(1-\beta_{xy}\right)t_{xy}^2-2\sqrt{\alpha_{xy}\beta_{xy}}t_{xy}+\left(1-\alpha_{xy}\right)=0.$$

Elementary analysis of the roots of this quadratic equation shows that it has at least one positive root if $(\{\alpha_{xy}>0,\beta_{xy}>0,\alpha_{xy}+\beta_{xy}\geq 1,\min\left(\alpha_{xy},\beta_{xy}\right)<1\})$. If these restrictions on the congestion function hold, there exists a pair of positive weights a_x and a_y that solves the quadratic equation, implying positive definiteness of the block element of the Hessian.

Note, however, that the scaling factors a_x and a_y are chosen for a pair of types x and y but cannot vary depending on which pair of types are considered. Therefore, the exact same value of t_{xy} must solve quadratic equations for all block elements of the Hessian independent of which pair of types (x, y) we consider. Therefore, for uniqueness we also require $\alpha_{xy}=\alpha_{xy'}=\alpha_{x'y}$, $\beta_{xy}=\beta_{xy'}=\beta_{x'y}$, $\frac{\varepsilon_{xy}}{\eta_{xy}}=\frac{\varepsilon_{xy'}}{\eta_{xy'}}=\frac{\varepsilon_{x'y}}{\eta_{x'y}}$ for all $x'\neq x$ and $y'\neq y$. □

A.3 Proof of Proposition 3

Proof. Step 1. First note that combining the two equilibrium conditions (18) and (19) one can derive

$$\frac{\mu p_l(h)}{1-\mu p_h(h)}=\frac{\left(1+\exp\left(\frac{1}{\theta}\left(\sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}}-u\right)\right)\right)}{\left(1+\exp\left(\frac{1}{\theta}\left(d-\sqrt{\frac{1-\mu p_h(h)}{\mu p_l(h)}}\right)\right)\right)}\frac{1}{\exp\left(\frac{1}{\theta}\left(\sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}}-u\right)\right)}. \tag{20}$$

Second, note that from equilibrium conditions it follows that $\mu p_h(h) > \frac{1}{2}$ holds iff $u > \sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}}$ and $\mu p_l(h) < \frac{1}{2}$ holds iff $\sqrt{\frac{1-\mu p_h(h)}{\mu p_l(h)}} < d$. Therefore, as long as $u > \sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}} > \frac{1}{d}$, the equilibrium is PAM. To check for consistency, taking the limit $\theta \to 0$ of (20), we conclude that $\mu p_h(h) \to 1$, $\mu p_l(h) \to 0$, $\frac{\mu p_l(h)}{1-\mu p_h(h)} \to \exp \frac{1}{\theta}\left(u - \sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}} + \sqrt{\frac{1-\mu p_h(h)}{\mu p_l(h)}} - d\right)$. For this expression to satisfy $u > \sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}} > \frac{1}{d}$, it must be that $u - \sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}} + \sqrt{\frac{1-\mu p_h(h)}{\mu p_l(h)}} - d = 0$; therefore,

$$\sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}} = \frac{1}{2}\left((u-d) + \sqrt{(u-d)^2 + 4}\right),$$

which indeed always satisfies the condition. This proves parts 3.1a and 3.2 of the proposition.

Step 2. Now assume vertical preferences, for instance $u > 1 > d$, without loss of generality. For there to be a mixing equilibrium, it must be that $\mu p_h(h) > \frac{1}{2}$ and $\mu p_l(h) > \frac{1}{2}$, which hold iff $u > \sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}}$ and $\sqrt{\frac{1-\mu p_h(h)}{\mu p_l(h)}} > d$. As we have already shown, under $u > \frac{1}{d}$ we obtain PAM; therefore, it must be that $\sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}} \leq u \leq \frac{1}{d}$. Taking the limit $\theta \to 0$ of (20), we conclude that $\mu p_h(h) \to 1$, $\mu p_l(h) \to 1$, $\frac{\mu p_l(h)}{1-\mu p_h(h)} \to \exp \frac{1}{\theta}\left(u - \sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}}\right) \to \infty$, unless $\sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}} = u$. In this case, $\frac{\mu p_l(h)}{1-\mu p_h(h)} \to \exp \frac{1}{\theta}\left(\sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}} - u\right) + 1 \to 2$, $\mu p_h(h) \to 1 - \frac{1}{u^2}$, $\mu p_l(h) \to 1$. For the limiting equilibrium to be mixing, it must be that $\sqrt{2} \leq u$, which completes the proof of 3.1b.

Step 3. For the equilibrium to exhibit NAM, it must be that $\mu p_h(h) < \frac{1}{2}$ and $\mu p_l(h) > \frac{1}{2}$ which hold iff $u < \sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}}$ and $\sqrt{\frac{1-\mu p_h(h)}{\mu p_l(h)}} > d$. This implies $u < \sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}} < \frac{1}{d}$. Taking the limit $\theta \to 0$ of (20), we conclude that $\mu p_h(h) \to 0$, $\mu p_l(h) \to 1$, $\frac{\mu p_l(h)}{1-\mu p_h(h)} \to 1$, unless $\sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}} = u$. In this case,

$$\frac{\mu p_l(h)}{1-\mu p_h(h)} \to \left(\exp \frac{1}{\theta}\left(u - \sqrt{\frac{\mu p_l(h)}{1-\mu p_h(h)}}\right) + 1\right) \Big/ \left(\exp \frac{1}{\theta}\left(d - \sqrt{\frac{1-\mu p_h(h)}{\mu p_l(h)}}\right) + 1\right) \to 2.$$

Since $\mu p_l(h) \to 1$ still, $\mu p_h(h) \to 1 - \frac{1}{u^2}$. Therefore, we obtain NAM as long as $u < \sqrt{2}$. This completes the proof of 3.1c and exhausts the possible relations between d and u.

Step 4. Now, take the opposite limit $\theta \to \infty$. From equilibrium conditions it follows that $\mu p_h(h) \to \frac{1}{2}$, $\mu p_l(h) \to \frac{1}{2}$, $\frac{\mu p_l(h)}{1-\mu p_h(h)} \to 1$. If we assume horizontal preferences,

$u > 1, \quad \frac{1}{d} < 1$, then it immediately follows that $\mu p_h(h) > \frac{1}{2}, \quad \mu p_l(h) < \frac{1}{2}$, which implies a PAM equilibrium. If instead we assume vertical preferences, $u > 1, \quad \frac{1}{d} > 1$, then similarly it immediately follows that $\mu p_h(h) > \frac{1}{2}$, $\mu p_l(h) > \frac{1}{2}$, which is a mixing equilibrium. This completes the proof of 3.3. $\square$

Appendix B: Identification in the 2×2 case

Consider again the case where there are $F = 2$ types of females (high and low) and $M = 2$ types of males (high and low); μ_h, μ_l are the number of females of each type; and v_h, v_l are the numbers of males of each type. Consider a general surplus $\Phi_{xy} = 2\begin{bmatrix} \Phi_{hh} & \Phi_{hl} \\ \Phi_{lh} & \Phi_{ll} \end{bmatrix}$. Following results from Section 2.5, assume that upon matching the surplus is split evenly between males and females and that the congestion function is $\phi_{xy} = C/\sqrt{p_x q_y}$ to guarantee both uniqueness and efficiency of equilibria. Assume that search costs are linear in effort and that the marginal costs are the same for all agents, denoted θ. The unique efficient equilibrium of this model is fully characterized by the set of equilibrium conditions as in Proposition 1. The matching rates are defined as follows:

$$m_{hh} = C\sqrt{\mu_h v_h}\sqrt{p_h(h)q_h(h)\mu_h v_h},\ m_{hl} = C\sqrt{\mu_h v_l}\sqrt{(1 - v_h p_h(h))\, q_l(h)\mu_h}$$
$$m_{lh} = C\sqrt{\mu_l v_h}\sqrt{p_l(h)\,(1 - \mu_h q_h(h) q_h(h))\, v_h}, m_{ll} = C\sqrt{\mu_l v_l}\sqrt{(1 - v_h p_l(h))\,(1 - \mu_h q_l(h))}.$$

The "identification" question is if we observe matching rates m_{xy} and quantities of unmatched agents, μ_x, v_y, what can we infer about Φ_{xy}? In the 2×2 case we can do this step by step. First, consider whether we can infer p_x and q_y. Invert the system of equations as follows:

$$\mu_h q_h(h) = \frac{m_{hh}^2/\left(C^2 \mu_h v_h\right)}{v_h p_h(h)}$$
$$v_h p_l(h) = \frac{m_{lh}^2/\left(C^2 \mu_l v_h\right)}{1-\mu_h q_h(h)} = \frac{m_{lh}^2/\left(C^2 \mu_l v_h\right)}{v_h p_h(h) - m_{hh}^2/(C^2 \mu_h v_h)}$$
$$\mu_h q_l(h) = 1 - \frac{m_{ll}^2/\left(C^2 \mu_l v_l\right)}{1-v_h p_l(h)} = 1 - \frac{\left(v_h p_h(h) - m_{hh}^2/\left(C^2 \mu_h v_h\right)\right) m_{ll}^2/\left(C^2 \mu_l v_l\right)}{v_h p_h(h) - m_{hh}^2/(C^2 \mu_h v_h) - m_{lh}^2/(C^2 \mu_l v_h)}.$$

Therefore, we obtain an equation with respect to p_1:

$$m_{hl}^2/\left(C^2 \mu_h v_l\right) = (1 - v_h p_h(h)) \left(1 - \frac{\left(v_h p_h(h) - m_{hh}^2/\left(C^2 \mu_h v_h\right)\right) m_{ll}^2/\left(C^2 \mu_l v_l\right)}{v_h p_h(h) - m_{hh}^2/(C^2 \mu_h v_h) - m_{lh}^2/(C^2 \mu_l v_h)}\right)$$

This can be simplified to a quadratic equation with respect to $v_h p_h(h)$ (changing notation from $m_{ij}/\mu_i v_j$ by m_{ij}):

$x^2C^2\left(C^2-m_{ll}^2-m_{lh}^2\right)-\left(C^2\left(C^2+m_{hh}^2-m_{hl}^2-m_{lh}^2-m_{ll}^2\right)+m_{hl}^2m_{lh}^2-m_{ll}^2m_{hh}^2\right)x+$
$m_{hh}^2\left(C^2-m_{ll}^2-m_{hl}^2\right)=0.$

This equation has two roots of the form

$v_hp_h(h)=\frac{\left(C^2\left(C^2+m_{hh}^2-m_{hl}^2-m_{lh}^2-m_{ll}^2\right)+m_{hl}^2m_{lh}^2-m_{ll}^2m_{hh}^2\right)\pm\sqrt{D}}{2\left(C^2-m_{ll}^2-m_{lh}^2\right)C^2}$, where

$D=\left(C^2\left(C^2+m_{hh}^2-m_{hl}^2-m_{lh}^2-m_{ll}^2\right)+m_{hl}^2m_{lh}^2-m_{ll}^2m_{hh}^2\right)^2$
$-4\left(C^2-m_{ll}^2-m_{hl}^2\right)\left(C^2-m_{ll}^2-m_{lh}^2\right)m_{hh}^2C^2.$

Thus, upon knowing the matching rates and the scaling factor, we can compute all the strategies:

$$\mu_hq_h(h)=\frac{m_{hh}^2/C^2}{v_hp_h(h)},\qquad v_hp_l(h)=\frac{m_{lh}^2/C^2}{1-\frac{m_{hh}^2/C^2}{v_hp_h(h)}},\qquad 1-\mu_hq_l(h)=\frac{m_{ll}^2/C^2}{1-v_hp_l(h)}.$$

This gives us a closed-form solution for the strategies conditional on matching rates.

The second step is to use the strategies p_x and q_y to infer the surplus Φ_{xy}. From the necessary conditions in proposition 1 we can write down the following relations (in log form):

$$\frac{\Phi_{hl}}{\theta}q_l(h)\frac{C}{\sqrt{\frac{(1-v_hp_h(h))}{v_l}q_l(h)}}=\frac{\Phi_{hh}}{\theta}q_h(h)\frac{C}{\sqrt{p_h(h)q_h(h)}}+\ln\frac{(1-v_hp_h(h))}{v_lp_h(h)}$$

$$\frac{\Phi_{ll}}{\theta}\frac{1-\mu_hq_l(h)}{\mu_l}\frac{C}{\sqrt{\frac{1-v_hp_l(h)}{v_l}\frac{1-\mu_hq_l(h)}{\mu_l}}}=\frac{\Phi_{lh}}{\theta}\frac{1-\mu_hq_h(h)}{\mu_l}\frac{C}{\sqrt{p_l(h)\frac{1-\mu_hq_h(h)}{\mu_l}}}+\ln\frac{1-v_hp_l(h)}{v_lp_l(h)}$$

$$\frac{\Phi_{lh}}{\theta}p_l(h)\frac{C}{\sqrt{p_l(h)\frac{1-\mu_hq_h(h)}{\mu_l}}}=\frac{\Phi_{hh}}{\theta}p_h(h)\frac{C}{\sqrt{p_h(h)q_h(h)}}+\ln\frac{1-\mu_hq_h(h)}{\mu_lq_h(h)}$$

$$\frac{\Phi_{ll}}{\theta}\frac{1-v_hp_l(h)}{v_l}\frac{C}{\sqrt{\frac{1-v_hp_l(h)}{v_l}\frac{1-\mu_hq_l(h)}{\mu_l}}}=\frac{\Phi_{hl}}{\theta}\frac{1-v_hp_h(h)}{v_l}\frac{C}{\sqrt{\frac{1-v_hp_h(h)}{v_l}q_l(h)}}+\ln\frac{1-\mu_hq_l(h)}{\mu_lq_l(h)}.$$

This is a system of four equations and four unknowns, and the solution can be obtained in closed form:

$$\frac{\Phi_{hh}}{\theta/C}=\frac{\sqrt{\frac{1-\mu_hq_l(h)}{1-v_hp_h(h)}}\left(\frac{1-v_hp_h(h)}{v_lq_l(h)}\ln\frac{1-v_hp_h(h)}{v_lp_h(h)}+\ln\frac{1-\mu_hq_l(h)}{\mu_lq_l(h)}\right)-\sqrt{\frac{1-v_hp_l(h)}{1-\mu_hq_l(h)}}\left(\frac{1-\mu_hq_h(h)}{\mu_lp_l(h)}\ln\frac{1-\mu_hq_h(h)}{\mu_lq_h(h)}+\ln\frac{1-v_hp_l(h)}{v_lp_l(h)}\right)}{\frac{1-\mu_hq_h(h)}{\mu_lp_l(h)}\sqrt{\frac{1-v_hp_l(h)}{1-\mu_hq_l(h)}\frac{p_h(h)}{q_h(h)}}-\frac{1-v_hp_h(h)}{v_lq_l(h)}\sqrt{\frac{1-\mu_hq_l(h)}{1-v_hp_l(h)}\frac{q_h(h)}{p_h(h)}}}$$

$$\frac{\Phi_{hl}}{\theta/C}=\sqrt{\frac{1-v_hp_h(h)}{v_lq_l(h)}}\left(\frac{\Phi_{hh}}{\theta/C}\sqrt{\frac{q_h(h)}{p_h(h)}}+\ln\frac{1-v_hp_h(h)}{v_lp_h(h)}\right)$$

$$\frac{\Phi_{lh}}{\theta/C}=\sqrt{\frac{1-\mu_hq_h(h)}{\mu_lp_l(h)}}\left(\frac{\Phi_{hh}}{\theta/C}\sqrt{\frac{p_h(h)}{q_h(h)}}+\ln\frac{1-\mu_hq_h(h)}{\mu_lq_h(h)}\right)$$

$$\frac{\Phi_{ll}}{\theta/C}=\sqrt{\frac{1-v_hp_l(h)}{1-\mu_hq_l(h)}}\left(\frac{\Phi_{lh}}{\theta/C}\sqrt{\frac{1-\mu_hq_h(h)}{\mu_lp_l(h)}}+\ln\frac{1-v_hp_l(h)}{v_lp_l(h)}\right).$$

Note that both stages of the closed-form solution depend in a non-trivial way on a scale factor C that is usually unknown and may be arbitrary. Simulations show that for a randomly generated set of matching rates there always exists a (tiny) interval of values of $C \in \left[\underline{C}, \overline{C}\right]$ for which all strategies take values in the unit interval $vp, \mu q \in [0,1]$ and all surplus values Φ_{xy} are real and non-negative. Simulation results show that as the scale factor moves along the interval, the implied surplus also scales from small values to large values, preserving the ordinal properties of the surplus matrix along the way but changing the proportions between elements somewhat.

This property provides an identification result: the "shape" of the payoff can be uniquely pinned down using only the matching rates. However, closed-form solutions can be derived only in the 2×2 case because the number of observed matching rates $(F * M)$ equals the number of unknown strategies $(F*(M-1) + M*(F-1))$ and equals the number of unobserved surplus values $(F * M)$. For a larger number of types, a closed-form solution is not possible, because there are more unknown strategies than there are observed matching rates. However, Monte-Carlo estimation for 3×3 and 4×4 cases, where estimating the surplus exactly from artificial data is computationally feasible in reasonable time, suggests that the same property of the mapping between matching rates and surpluses holds as for the 2×2 case. For any observed or artificially generated matching rate matrix, it is possible to find multiple surplus matrices (roughly proportional to each other and preserving ordinal properties between elements) that match the data exactly.

Appendix C: Calibration using contact rates

If contact rates are available from a dating website, for example, one can infer for each type of male and female what their distributions of interests—$q_y(x)$ and $p_x(y)$— are. Since these distributions sum to 1, the data contain observations with $2 \times M \times F - M - F$ degrees of freedom. Assuming non-transferable utility and knowledge of search costs, these data allow the researcher to identify the shape of the payoff functions for each type of male and female, ε_{xy} and η_{xy}, respectively, which have a total of $2 \times M \times F$ independent parameters.

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 (9-10) we obtain

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

$$\ln\frac{q_y^*(x)}{q_y^*(x')} = \frac{\eta_{xy}}{\theta_y} p_x^*(y)\,\phi_{xy}^* - \frac{\eta_{x'y}}{\theta_y} p_{x'}^*(y)\,\phi_{x'y}^*. \tag{21}$$

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.

Appendix D: Empirical Application

D.1 Sorting

Figure 4 displays sorting by income. We divide people into deciles, except for the first two (5 and 15), which we bunch into one denoted by "5," as these include only individuals with zero income. From the payoff function we can see that the productive motive would drive everyone to aim at someone in the highest quintile. However, when we look at the strategies, females do primarily target males in the highest decile but males randomize primarily among the highest four deciles and still give some non-negligible weight to females below that. The number of matches reflects the combination of the strategies of males and females.

For this case, the HI=0.06 and the AI=0.19. This means that preferences are clearly vertical and the degree of assortativeness indicates a mixing equilibrium. If we recover preferences using the Choo and Siow model, we would find HICS=0.56, which would point toward preferences that are much more horizontal than those produced by our model.

The results of sorting by age are shown in Figure 5. We take people between the ages of 21 and 40 and sort them into 20 bins. From the payoff function we can see that males prefer females between ages 27 and 31, while females prefer males between ages 29 and 33. However, because those males and females are unlikely to reciprocate interest, both males and females have to settle for partners closer to their own age,

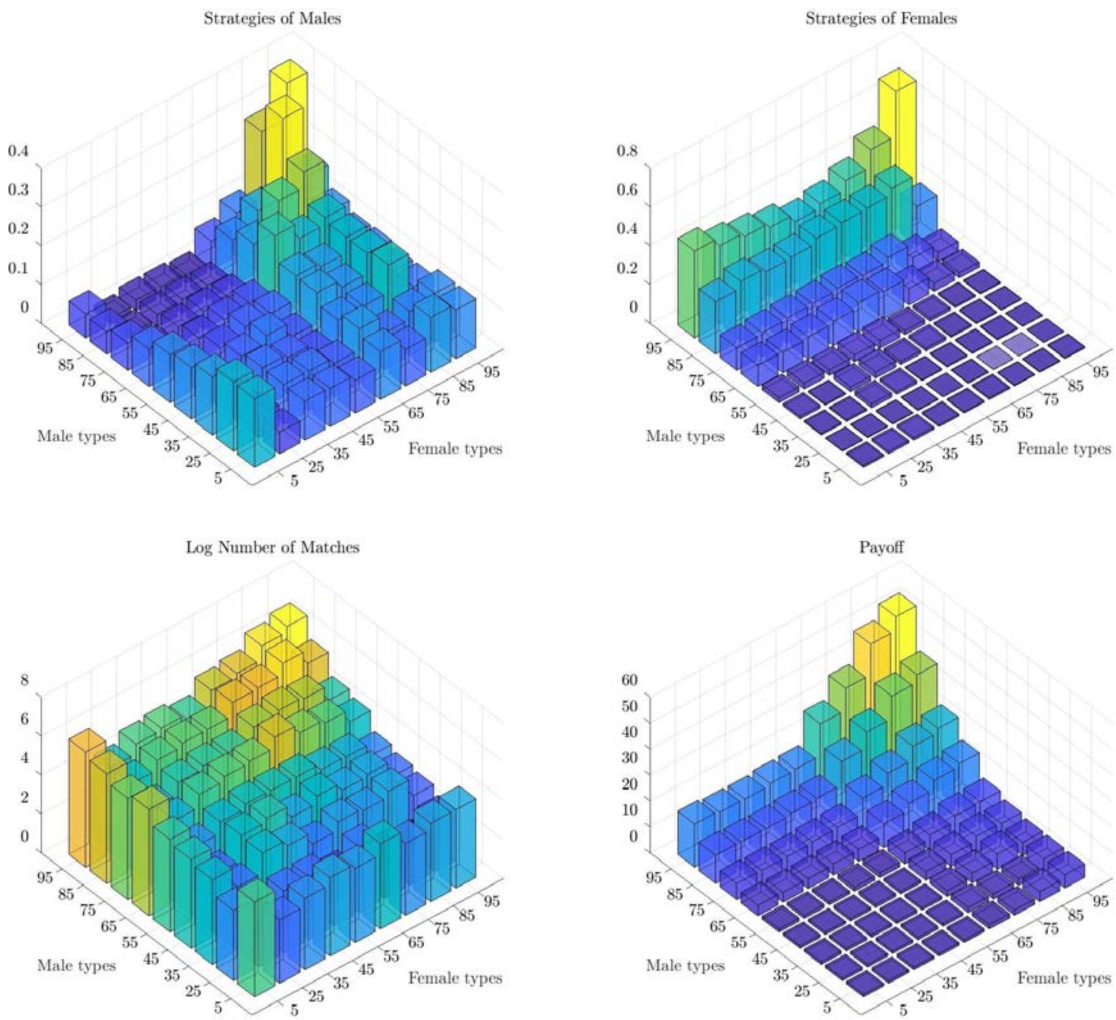

Figure 4: Sorting by Income

and those are the ones they end up targeting. This suggests an important role for the strategic motive in matching by age.

For this case, the HI=0.18, indicating preferences that are more vertical towards individuals between 27 and 31 for males and between 29 and 33 for females, and the AI=0.61, reflecting assortative matching with a little mixing. When we calculate the HICS we find it to be 0.74, which would be misleadingly indicative of horizontal preferences according to our model.

Intuitively, sorting by age is similar to the frictionless assortative matching story where the strategic motive is the main driving factor behind matching. This is a clear example of preferences by age exhibiting verticality: broadly consistent with the online dating evidence, but contrary to what a non-strategic model like Choo and Siow would conclude. This example highlights the ability of our model to identify vertical preferences even when the matching-rate pattern seems assortative, suggesting horizontal preferences.

Figure 6 characterizes sorting by education. We divide people into three bins. Those who have completed high school (S), those who have completed college (C), and those who have a postgraduate degree (P). From the payoff function we can see that everyone would be better off being with someone that has a postgraduate degree. However, both males and females with less than a postgraduate degree target individuals with a college degree. This indicates that those with a college degree settle for someone like themselves and that those with just a high school diploma aim higher by targeting someone with more education.

For this case both the HI and the AI equal zero, which is a clear indication of vertical preferences that generate a mixing equilibrium, reflecting very high importance of the strategic motive. If we were to take a model of unobserved characteristics like in Choo and Siow (2006) and use it to recover the underlying surplus and then calculate our HI, we would find HICS to be 1, which would indicate horizontal preferences when according to our model they really are vertical.

Figure 7 shows sorting by race. We divide people into four bins: white (Wh), Hispanic (Hi), black (Bl), and Asian (As). From the payoff function, we can see that

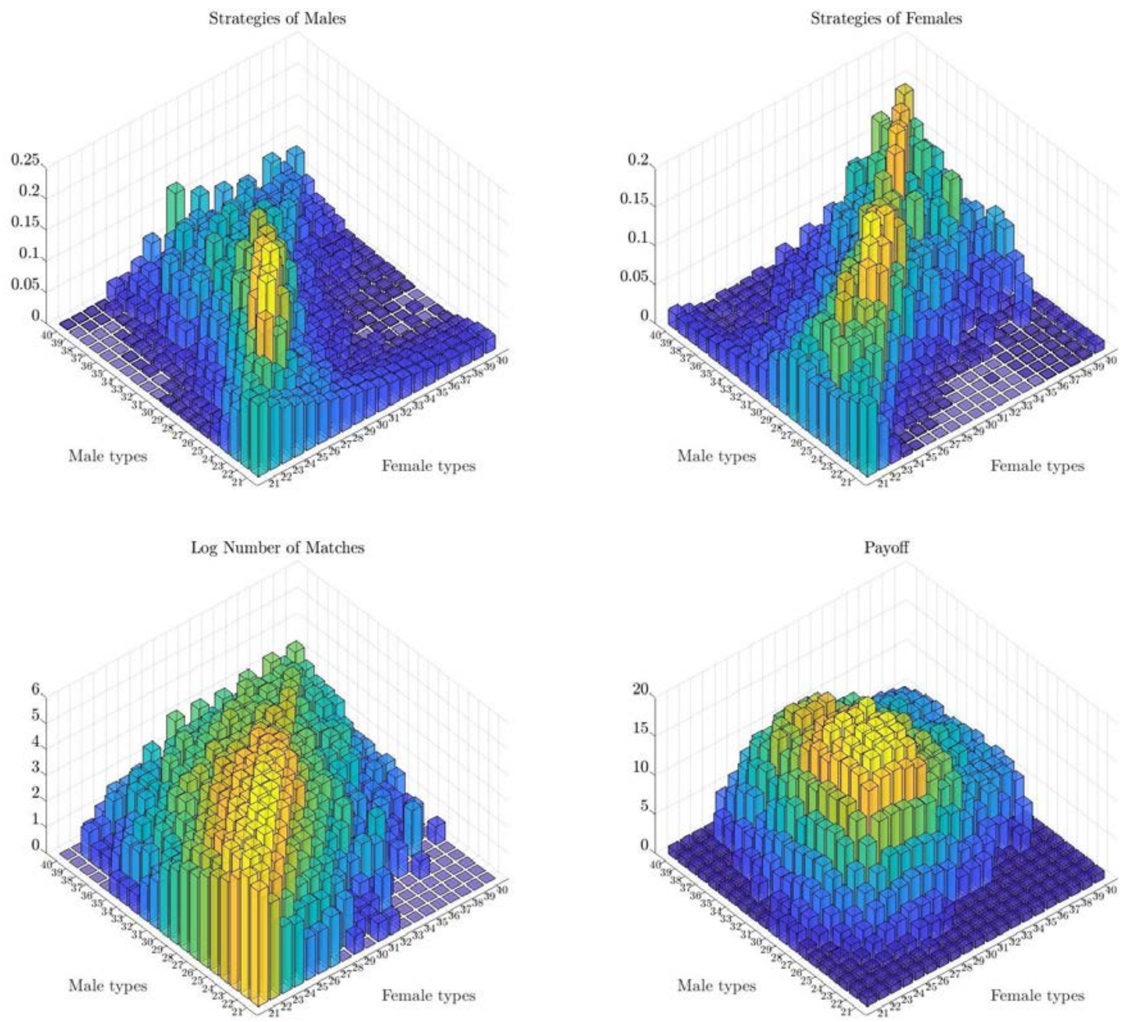

Figure 5: Sorting by Age

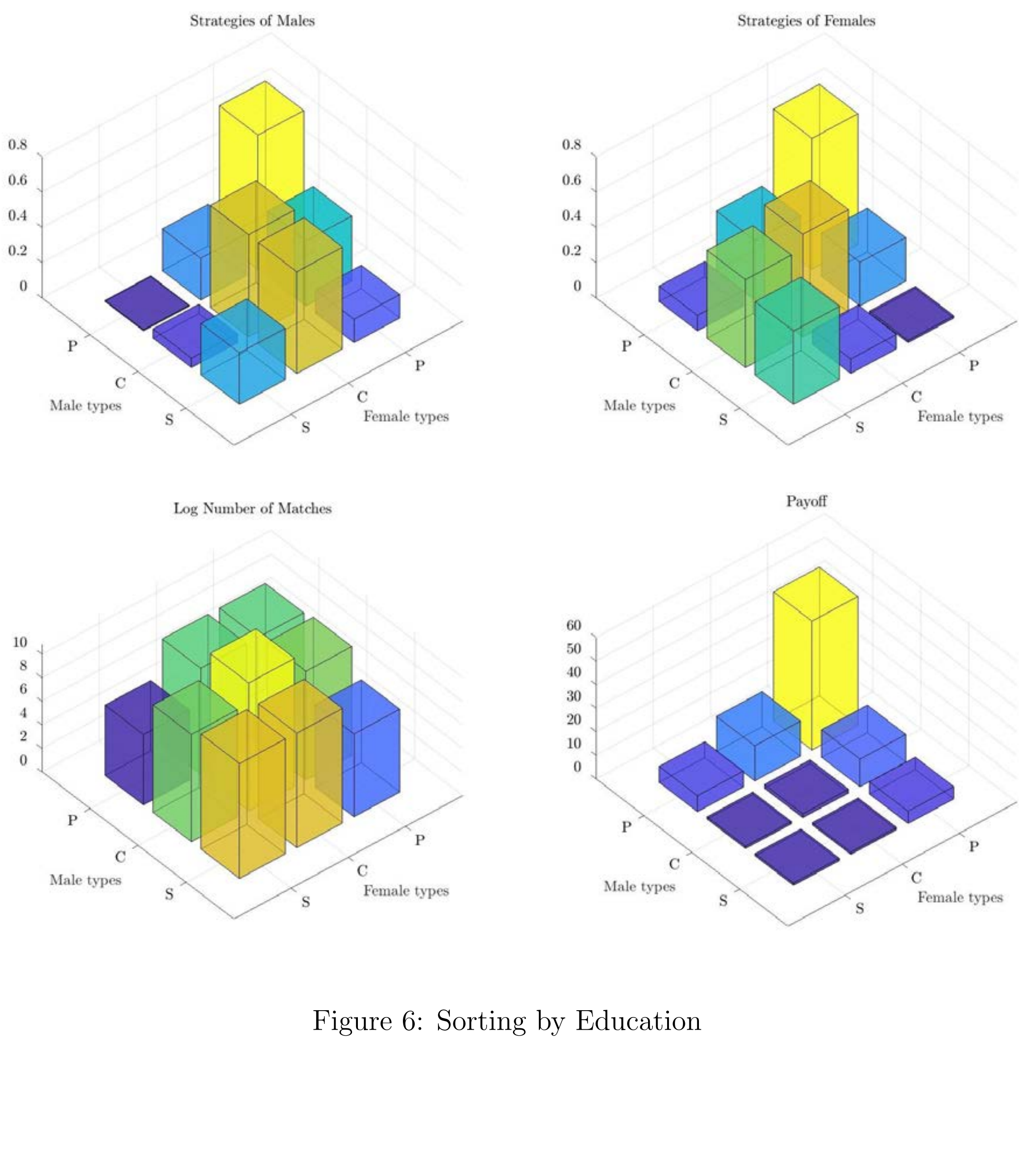

Figure 6: Sorting by Education

everyone has a preference for a white partner, closely followed by an Asian partner. However, from the strategies, we see that with the exception of black males and females, everyone targets someone of their same race.

For this case, HI=0.33, pointing toward preferences that are more vertical, and AI=1. This is another example that highlights the ability of our model to identify vertical preferences even when the matching rate pattern seems assortative, deceptively suggesting horizontal preferences.

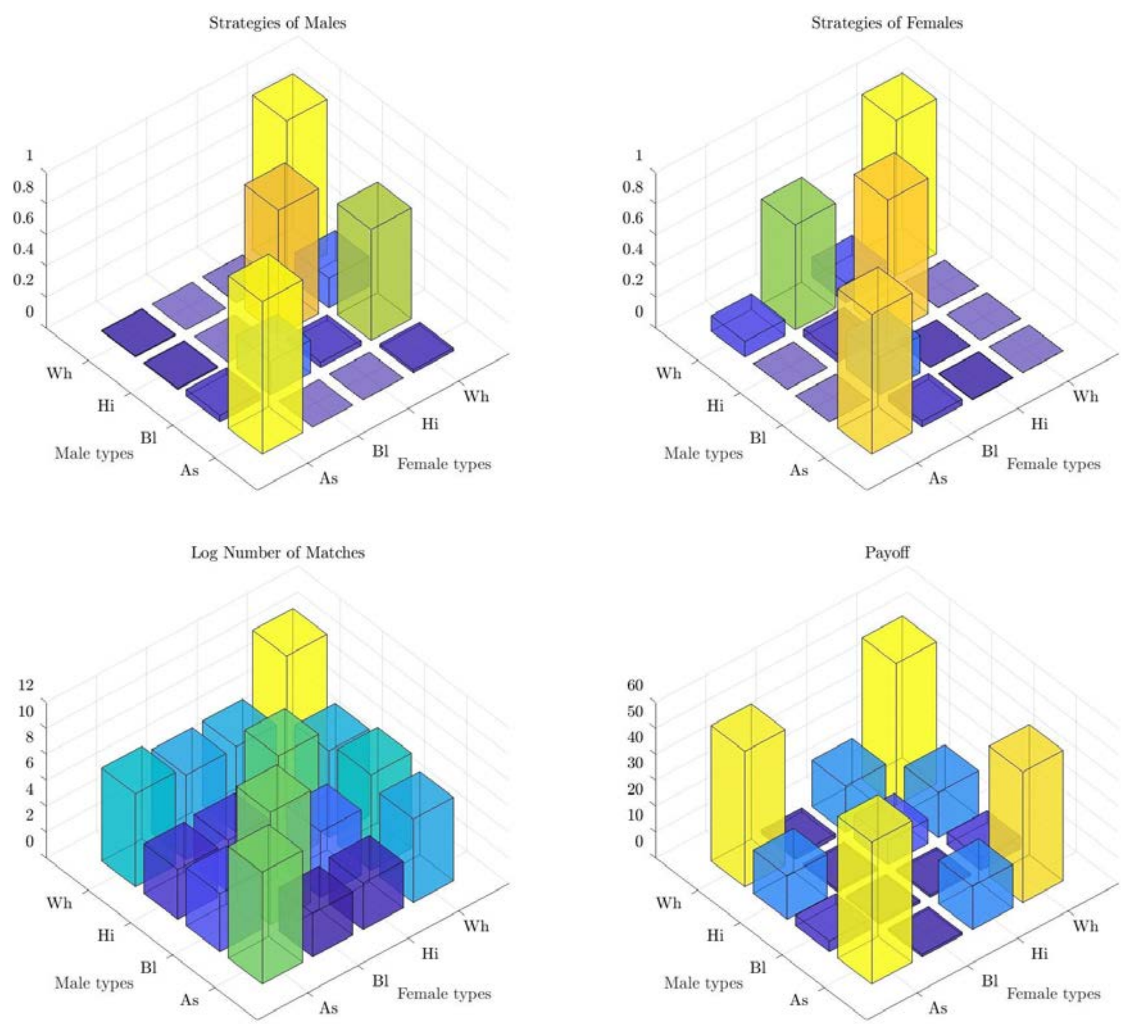

Figure 7: Sorting by Race

D.2 Desirability Indices

The middle panels plot the main target of each male and female type; in other words, this panels show to which type males and females assign the highest probability. The left panels show the DI of males from the perspective of females, and the right panels show the DI of females from the perspective of males.

Figure 8 shows the results by income. Both males and females with higher income are more desirable, and desirability falls as income goes down to the 45th decile for males and to the 35th decile for females and increases only slightly for the lowest deciles.

Figure 9 shows the results by age. We can see that women older than 33 target males of their same age and younger women target slightly older males, while males older than 31 target slightly younger women and males younger than 29 target females of their same age, with the exception of 21-year-olds. We find that between 21 and roughly 32, desirability increases with age for males and then decreases with age. For females it is very similar; their desirability increases with age between 21 and 31 and then decreases with age. This result makes sense because being a single male or female above a certain age might be considered a "red flag," hence the non-monotonicity.

Figure 10 shows the results for education. We can see that both males and females pursue partners out of their league, everyone targets individuals with a postgraduate degree. Also, the DIs generated by the model show that both males and females with a postgraduate degree are more desirable than those with a college degree and more desirable than those with just a high school diploma.

Finally, Figure 11 shows the results by race. We can see that for both males and females, whites are preferred, followed by Asians, Hispanics, and blacks, but every type's main target is of their own race.

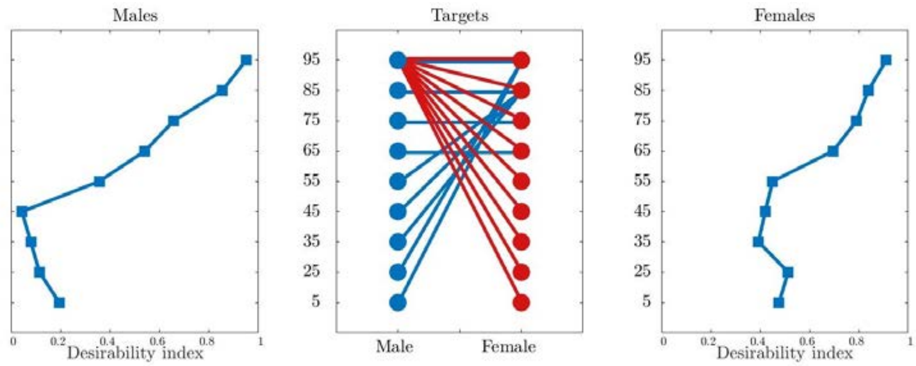

Figure 8: Desirability Index by Income

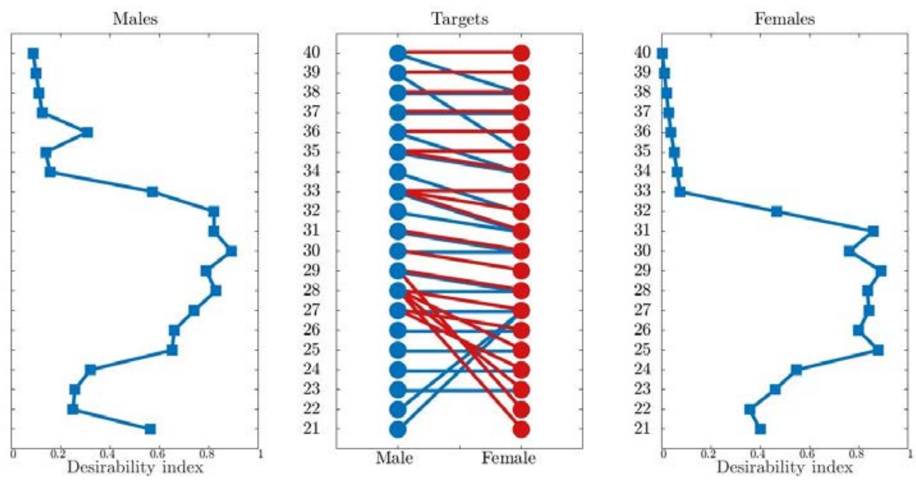

Figure 9: Desirability Index by Age

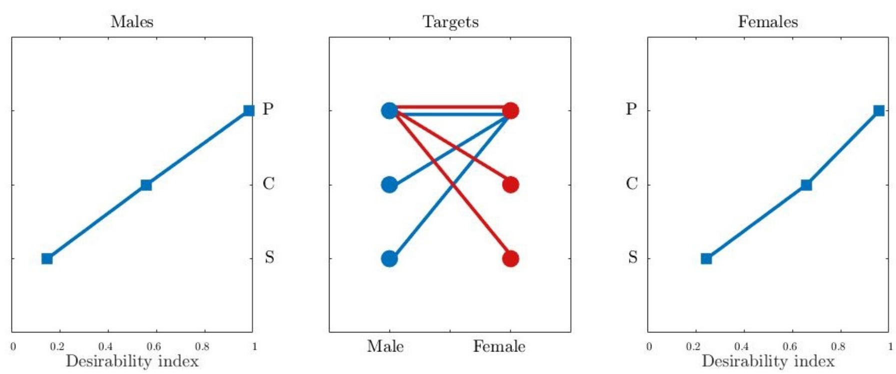

Figure 10: Desirability Index by Education

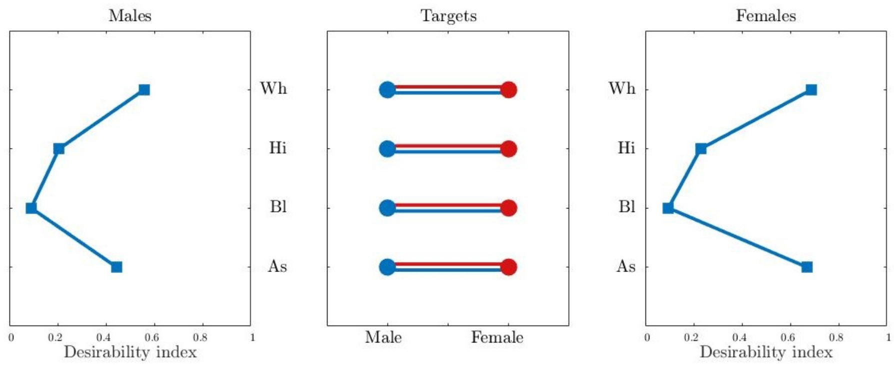

Figure 11: Desirability Index by Race

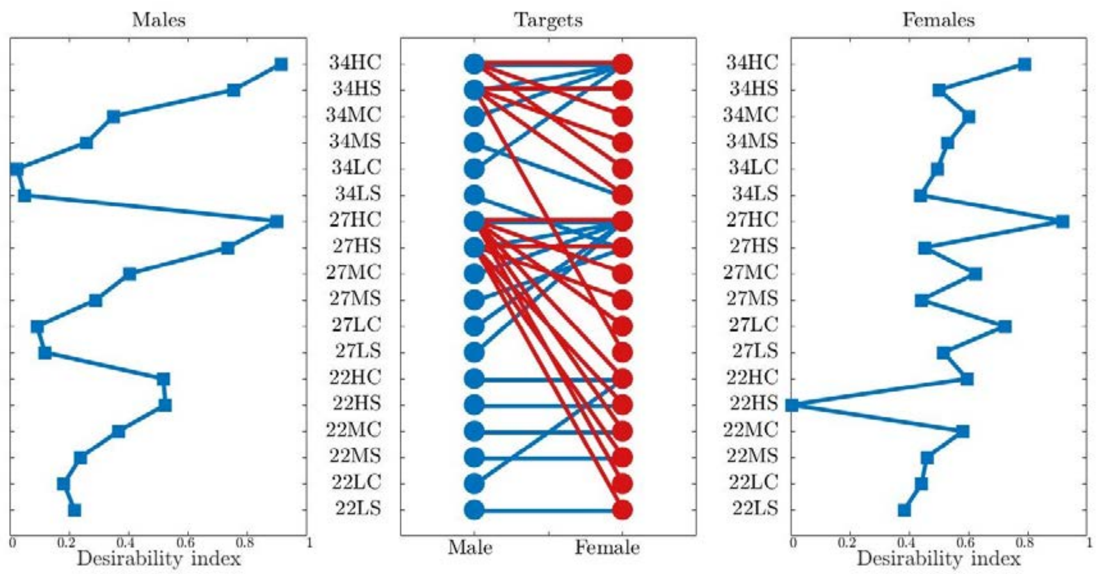

Figure 12: Desirability Index Sorting by Multidimensional Characteristics

The study of matching with multidimensional attributes quickly runs into the empirical problem that many pairs of types produce zero matches in equilibrium. While our model is able to match such patterns reasonably well, the estimates imply a substantial payoff differential between acceptable matches and undesirable matches. This differential needs to be large enough to incentivize men and women to identify the undesirable types and exclude them from their pool of potential contacts to avoid meeting them. However, in reality a different mechanism is likely at work in such cases. Both men and women likely meet those undesirable types but reject them at the match formation stage and decide to simply wait to meet a better type. To explain these cases in a satisfactory way, in addition to cognitive costs of distinguishing between types we need to introduce time costs into our model. To achieve that goal, in appendix E, building on Adachi (2003) we develop a model of repeated interactions where agents both choose whom to meet under a cognitive constraint and then form an "exclusion set" rejecting some of the matches upon meeting. This mechanism not only allows for a better rationalization of absence of matches between some of the types, but also affects the overall sorting pattern. We leave the task of estimating the repeated model on aggregate data for future research.

Appendix E: Repeated two-sided model

Here we extend the two-sided matching model to a repeated setting. Following Eeckhout (1999) and Adachi (2003), we assume that at the moment that a male of type y and a female of type 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 $\rho = 1$ represents a frictionless case, which implies that agents could wait for their preferred match indefinitely at no time cost to them. 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 distinguish among potential partners, i.e., the parameter θ approaches zero, the model, possibly using a refinement permitting only stable matchings, also reproduces the frictionless matching outcome.

We denote by v_x the continuation value of female x and by 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-zero problems of the agents are like before:

$$\max_{p_x(y)} Y_x = \sum_{y=1}^{M} \mu_y EU_x(y)\, Q_y(x)\, p_x(y) - \theta_x \left(\sum_{y=1}^{M} \mu_y p_x(y) \ln \left(p_x(y) \sum_{y=1}^{M} \mu_y \right) \right),$$

$$\max_{q_y(x)} Y_y = \sum_{x=1}^{F} \mu_x EU_y(x)\, P_x(y)\, q_y(x) - \theta_y \left(\sum_{x=1}^{F} \mu_x q_y(x) \ln \left(q_y(x) \sum_{x=1}^{F} \mu_x \right) \right).$$

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

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

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

And the expected utilities from meeting are either equal to match utilities if both partners agree to a match or equal to continuation values if they do not:

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

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

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 (NTU), but it can easily be extended to the case of transferable utility (TU). Specifically, the TU case allows for redistributing the surplus in the cases when joint surplus of the match exceeds the sum of the continuation values of the agents. Therefore, the last two equations are replaced in the TU case by

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

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

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 because all continuation values are zero. In Figure 13, using a simple payoff structure that exhibits vertical preferences for three males and three females and no congestion ($\phi = 1$), we illustrate 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 three pairs represents one-to-one matching, while the case with nine pairs implies that all possible pairings are observed. Like in the one-shot model, there is a threshold level of cognitive costs that generates multiplicity of equilibria. There are also small islands of multiplicity generated by the same mechanism as in the Adachi model. In the region where both costs are relatively high, all pairs of types are matched with some frequency in the unique equilibrium.

In Figure 14, for our 2×2 symmetric example from Section 2.6 with parameter $u = 1.3$ and $d = 0.5$, we illustrate the regions of the parameter space (θ, ρ) in which the equilibrium is non-unique (red) and has different sorting patterns. Under low patience, equilibrium is NAM for low values of search costs and mixing for high values of search costs, as would be predicted by Proposition 3. For high patience the equilibrium has a PAM structure due to agents choosing to reject subpar potential matches upon meeting them and instead wait to meet their best matches. For intermediate values of patience, there is a region of multiplicity of equilibria, where for each combination of parameters multiple rejection/acceptance patterns can be supported by continuation values as equilibria.

The last example highlights that, even though under the conditions that guarantee uniqueness of equilibria in the one-shot model the repeated model can have multiple equilibria, the repeated model could nevertheless be successfully estimated using the same methodology. This is because all cases of multiple equilibria are characterized by different rejection/acceptance patterns in those equilibria. The empirically observed pattern would always uniquely pin down the relevant equilibrium in the case of multiplicity. Note, however, that another potential source of multiplicity emphasized in the literature can be due to the endogeneity of the stationary distribution of types as in Shimer and Smith (2000).

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 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 the parameter θ, which describes an agent's 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

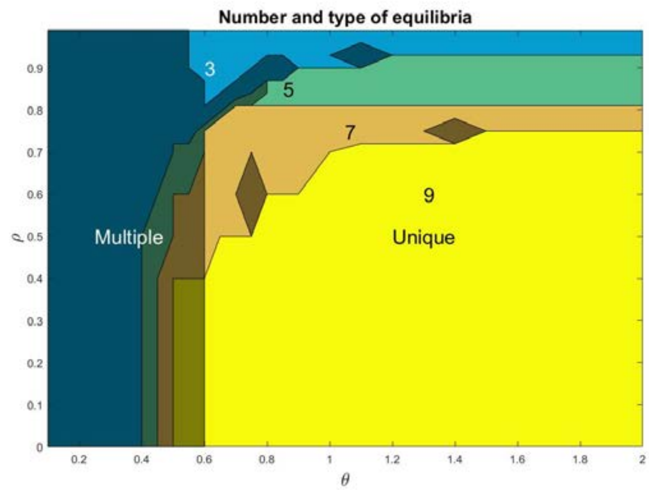

Figure 13: Number and Types of Equilibria Depending on Parameters, no Congestion

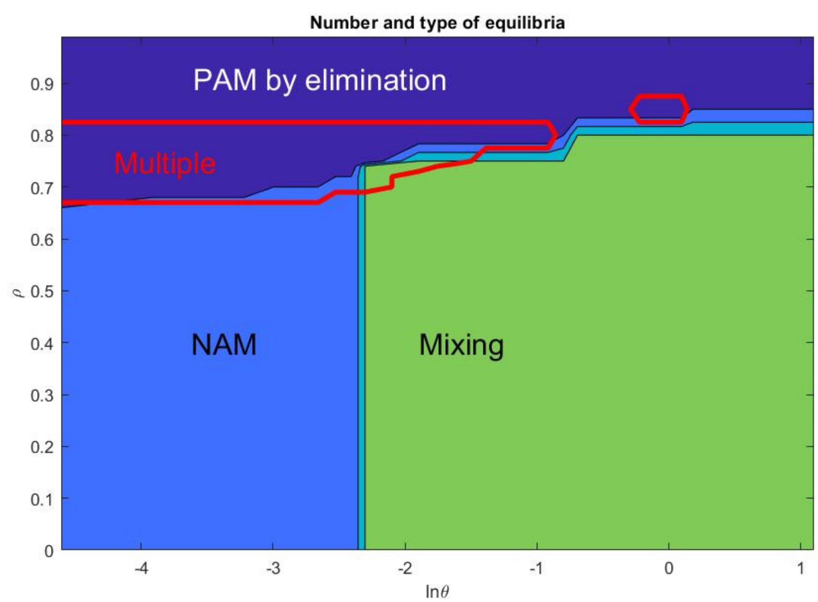

Figure 14: Number and Types of Equilibria Depending on Parameters, Congestion

constrained by the physical costs. Ours is an explicit model of how agents choose whom to meet. Second, the repeated model makes clear the source of the difference between the TU and the NTU cases. If agents are able to reject potential partners deemed not good enough, then it is important to know 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.