Research Division
Federal Reserve Bank of St. Louis
Working Paper Series

Targeted Search in Matching Markets

Anton Cheremukin
Paulina Restrepo-Echvarria
and
Antonella Tutino

Working Paper 2014-035B
http://research.stlouisfed.org/wp/2014/2014-035.pdf

May 2016

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to Federal Reserve Bank of St. Louis Working Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors.

Targeted Search in Matching Markets*

Anton Cheremukhin, Paulina Restrepo-Echavarria, Antonella Tutino

May 20, 2016

Abstract

We endogenize the degree of randomness in the matching process by proposing a model where agents have to pay a search cost to locate potential matches more accurately. The model features a tension between an agent's desire to find a more productive match and to maximize the odds of finding a match. This tension drives a wedge between the shape of sorting patterns and the shape of the underlying match payoff function. We show the empirical relevance of the latter prediction by applying the model to the U.S. marriage market.

JEL: E24, J64, C78, D83.

Keywords: Matching, sorting, assignment, search.

*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. Authors thank Fernando Alvarez, Gady Barlevy, Hector Chade, Jan Eeckhout, Mike Golosov, Christian Hellwig, Oleg Itskhoki, Boyan Jovanovic, Ricardo Lagos, Jim Peck, Robert Ulbricht, Pierre-Olivier Weill, and David Wiczer for discussions and suggestions which helped improve the paper. We also thank seminar participants at multiple conferences. All 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

The search for a partner—whether in business or in personal life—includes both productive and strategic considerations: People seek productive partnerships that maximize their pay-offs, but also search strategically by considering the likelihood a potential partner will also be interested in them. As a result, we observe that matching is not perfectly assortative nor purely random.

In this paper we propose a model that endogenizes the degree of randomness in the search and matching process by formalizing the fact that when looking for a match agents have two incentives: first, maximize their payoffs (which we refer to as the productive motive), and second, maximize the odds of forming a match (which we refer to as the strategic motive).

Our model can be empirically tested and we use it to rationalize observed matching rates in the U.S. marriage market and estimate the underlying match payoff function between males and females. We show that the tension between the productive and the strategic motive drives a wedge between the shape of sorting patterns and the shape of the underlying match payoff function recovered by the model. This result has important implications for empirical inference as it implies that one can derive the wrong conclusion about the form of productive complementarities, and hence, the amount of losses from mismatch in the marriage market by just looking at the matching rate.

We build on the frictionless matching environment of Becker (1973) with two-sided heterogeneity, and assume that even though both males and females know the distribution and their preferences over types of the other side of the market, there is noise and they cannot locate potential partners with certainty.

Males and females can pay a search cost to locate prospective matches more accurately, however accurately locating their best match is too costly so agents do not target one person with certainty. Instead, each male and female optimally chooses a discrete probability distribution over all types. An element of this distribution reflects the likelihood with which they will target each potential match, based on its expected pay-off. A more targeted search (placing a higher probability on specific types) is costlier than choosing a more spread out distribution across types because it implies a bigger search effort and cost, but allows males and females to target their better matches more accurately and guarantees a higher expected payoff.

Once each male and female select their optimal probability distribution, they draw a match from that distribution. If the outcome of the draw is reciprocated and is mutually beneficial, a match is formed and the output of the match is split between the two parties. Because the search involves balancing costs and benefits about prospective matches, agents will optimally choose partially targeted strategies—they will not place a unit mass on a particular potential partner—and some participants will remain unmatched.

Both the productive and the strategic motive are at play when choosing the optimal distribution. The productive motive drives agents to target the person that renders a higher payoff, while the strategic motive drives agents to target the person that is more likely to be interested in them.

These two motives play a crucial role in shaping two theoretical predictions of our model: (1) the uniqueness and inefficiency of the equilibrium; (2) the implications for sorting patterns.

In terms of the uniqueness of the equilibrium, the strategic motive makes the search strategies for both sides of the market complementary; both sides will be more interested in someone who is more likely to be able to locate

them (reciprocate their interest). The strength of these complementarities is key for determining the uniqueness of the equilibrium.[1] In the optimal assignment model, complementarities between search strategies are strong and lead to multiplicity of equilibria,[2] whereas complementarities are absent in random matching models. We show that an increase in the search cost makes the search strategies of market participants less complementary and eliminates the multiplicity of equilibria.

Whether the productive or the strategic motive dominates depends on the search cost. When search costs are sufficiently high, the productive motive dominates and there is always a positive probability for a low type to be matched with a high type (because the low-type knows that the high-type cannot locate another high-type with accuracy). In this case the equilibrium that emerges is unique and of the mixing type, i.e., some high types match with high types and some high types match with low types. If the costs are below a certain threshold, the strategic motive dominates and there are multiple equilibria that can deliver positive assortative matching, negative assortative matching, or a mixing equilibrium.[3]

As a result, the equilibria of our model lie *in between* the random matching outcome and the frictionless assignment outcome and encompass both outcomes as limiting cases. In the extreme, if search costs tend to infinity, then agents will choose not to incur the cost and the optimal strategy will correspond to the random matching outcome. If search costs are zero, then

[1]The complementarities in our model are different from search externalities studied by Diamond (1982); where an increase in the number of participants makes it easier for one side of the market and more difficult for the other side to find a match. In our model, the complementarity arises because if one person targets another, the other has an incentive to reciprocate.

[2]Many models of directed search, e.g. Shimer (2005), introduce additional assumptions in order to select one of those equilibria. The requirement of "stability" is another common equilibrium selection strategy.

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

agents can locate their best match with certainty and the outcome is a frictionless assignment. Furthermore, we show that the equilibria that emerge from a positive and finite cost are inefficient relative to the constrained Pareto allocation, although the outcomes are constrained efficient in both limiting cases.

The intuition behind the inefficiency is as follows: When an agent decides to pay a higher cost, she is choosing a more targeted strategy. This increases the probability of ending up with a more productive match and lowers the probability of being paired with a less productive one. Because of the complementarities in our model, high types are more likely to assign a higher probability in their distribution to other high types. Low types would then assign a lower probability to this (high-type) participant in their distribution and target someone else. Market participants are unable to fully internalize this positive externality. Males and females fail to appropriate all social benefits of their actions, and, as a result, the quantity and the quality of matches are both inefficiently low.

The second theoretical prediction of our model is that the shape of the payoff function determines whether the productive and the strategic motives are aligned or lead in different directions. Furthermore, the correlation between the shape of the equilibrium matching rate and the shape of the underlying payoff function depends on whether these motives co-move or clash. If the payoff function exhibits relative advantage,[4] the productive motive points all agents in different directions and the strategic motive ensures that the agent that renders the higher payoff is also more likely to reciprocate (because there is no competition for the same agent). Hence, the resulting

[4]We say that the payoff exhibits relative advantage when (1) a high-type female is better off with a high-type male and a low-type female is better off with a low-type male and vice versa or (2) more generally when for each type on one side the best option on the other side is different.

correlation between the matching rate and the underlying payoff function is high. However, if the payoff function exhibits absolute advantage,[5] the productive motive points all agents in the same direction, and the strategic motive tends to coordinate agents' to place a higher probability on those whom their competitors are less likely to consider (to improve the odds of forming a match). Hence, the resulting correlation between the matching rate and the underlying payoff function can be low.

To show the empirical relevance of this result, we take our model to the data. We use a standard dataset for matching in the U.S. marriage market, and construct matching rates based on income, age, and education separately. Our theory is testable in the sense that it places certain restrictions on equilibrium outcomes which may or may not be rejected by the data. We provide a simple example showing that there exist matching rates that cannot be explained by our model. Nevertheless, we find that the model does a very good job rationalizing the observed marriage matching rates based on income, age, and education, and for these three cases we estimate the underlying payoffs implied by our model.

We find that the payoffs based on income and age exhibit absolute advantage and that the payoff based on education exhibits both relative and absolute advantage. For income, this means that marrying someone with higher income is always better. For age, this means that females have a strong preference for older males independent of their own age, while males are virtually indifferent about the age of their spouse. For education, people with low levels of education and people with high levels of education prefer someone with their same level of education, generating a region of relative advantage. However, people with a medium level of education tend to prefer

[5] We say that the payoff exhibits absolute advantage if there is a single type on each side of the market that is preferred by everyone.

highly educated people, generating a region with absolute advantage.

The correlation between the three observed matching rates and the recovered payoff functions ranges from 0.4 to 0.7. This means that strategic considerations uncovered by endogenizing randomness can drive a significant wedge between the shape of the observed sorting pattern and the shape of the underlying payoff function. Ignoring these considerations may result in misleading implications about the degree of mismatch present in the market and hence about the size of the losses associated with it.

Our paper effectively blends two approaches to introducing randomness used in the empirical literature. The first approach introduces search frictions by assuming that it takes time to find a match, as in Shimer and Smith (2000). The second approach introduces unobserved characteristics as a tractable way of accounting for the deviations of the data from the stark predictions of the frictionless model, as in Choo and Siow (2006) and Galichon and Salanie (2012).

We build on the discrete choice rational inattention literature—i.e., Cheremukhin, Popova, and Tutino (2015) and Matejka and McKay (2015)—that endogenizes the multinomial logit discrete choice model by introducing cognitive constraints that capture limits to processing information. Consequently, 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 predicts strong interaction between distributions of randomness in matching that is driven entirely by conscious strategic choices of agents, rather than by some unobserved characteristics with fixed distributions.

The search and matching literature has models like Menzio (2007) and Lester (2010) that nest directed search and random matching to generate

outcomes with an intermediate degree of randomness. Our paper contributes to the literature by providing a theory that produces equilibrium outcomes featuring endogenous randomness in between random matching and the frictionless assignment, without nesting these two frameworks.

Finally, the paper contributes to the literature on directed search and coordination frictions, as in Eeckhout and Kircher (2010) and Shimer (2005).[6] The directed search paradigm generally predicts socially efficient and assortative equilibrium outcomes. In contrast, our targeted search model does not appear to possess a market mechanism that can implement the constrained efficient allocation, nor does it guarantee assortativeness.

The paper proceeds as follows: Section 2 describes the model and derives the theoretical predictions. In Section 3 we take the model to the U.S. marriage market data. Section 4 states some final remarks.

2 The Model

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

Agents face a search cost and choose search effort by choosing an optimal probability distribution over types. This distribution reflects the likelihood of locating a particular agent. A more targeted search, or a probability distribution that is more concentrated on a particular group of agents implies

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

a higher cost as it means that the agent is exerting more search effort to locate a potential match of a certain type. As agents optimally choose how targeted their search strategy is by choosing a probability distribution, the degree of randomness in matching is endogenous. When they play a more targeted strategy, the degree of randomness goes down.

The probability distribution needs to satisfy two properties: 1) By the nature of the choice between a finite number of options, the distribution must be discrete; and 2) for strategic motives to play a role, agents should be able to vary each element of the distribution and consider small deviations of each element in response to changes in the properties of the options. Hence, this probability distribution cannot be confined to a specific family of distributions.

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

After choosing their optimal probability distribution over types, both males and females make a draw from their distribution. If the draw is reciprocated, a match is formed if it is mutually beneficial and the output from the match is split between the two parties. As the search involves balancing the costs and benefits of prospective matches, some participants will not find partners immediately.

2.1 The environment

There are F females indexed by $x \in \{1, ..., F\}$ and M males indexed by $y \in \{1, ..., M\}$. Both males and females are heterogeneous in types and are actively searching for a match. A match between female x and male y generates a payoff Φ_{xy}.[7] 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 share of the payoff appropriated by the female ε_{xy} and the share of the payoff appropriated by the male η_{xy} such that $\Phi_{xy} = \varepsilon_{xy} + \eta_{xy}$.[8] The payoff and the split generated by any potential (x, y) match are known ex-ante to female x and male y.

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

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

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

[8] Given that our aim is to understand the implications of having endogenous randomness, we focus on the simplest model and assume fixed payoff shares.

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

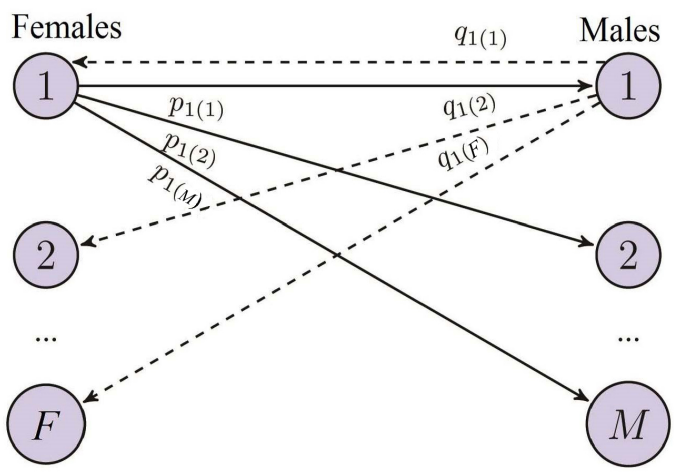

Figure 1. Strategies of males and females

Since negative payoffs lead to de facto zero payoffs due to the absence of a match, we can assume that all payoffs are non-negative:

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

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

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

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

The cost function is given by $c_x(\kappa_x) = \theta_x \kappa_x$, where θ_x is the search cost. Here, we are using the linear case for simplicity, but all of our proofs will hold for more general cost functions. As mentioned earlier, κ_x reflects search effort and needs to accommodate the full choice of a discrete distribution and attach a cost to it. One function that satisfies these requirements is the following:[9]

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

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

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

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

which allows them to target their better matches more accurately.

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

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

where

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

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

2.2 Matching Equilibrium

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

Theorem 1. *A matching equilibrium exists.*

Proof. Note that since the strategies are discrete, this is a normal-form game. The equilibrium of the matching model can be interpreted as a Nash equilibrium of this game. The set of distributions mapping compact sets into compact sets is a lattice under the natural ordering. Hence, all the results for lattices described by Vives (1990) apply to it. Since cross-derivatives of objective functions in our case are all non-negative, this game is supermodular. Hence, there exists a Nash equilibrium. □

Theorem 2. *The matching equilibrium is unique if*

a) cost functions are non-decreasing and convex;

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

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

Proof. The payoffs of all males and females are continuous in their strategies. They are also concave in these strategies when cost functions are (weakly) increasing and convex in κ. "Diagonal dominance" conditions (b) and (c) guarantee that the Hessian of the game is negative definite along the equilibrium path. Then, by the generalized Poincare-Hopf index theorem of Simsek, Ozdaglar, and Acemoglu (2007), the equilibrium is unique. □

Note that the assumptions we make to prove uniqueness are by no means restrictive. The assumption that cost functions are non-decreasing and convex is a natural one. The additional "diagonal dominance" conditions in our case can be interpreted as implying that the search cost of reducing noise should be sufficiently high for the equilibrium to be unique. If these conditions do not hold, then there can be multiple equilibria. This is a well-known outcome of the assignment model, which is a special case of our model under zero search costs. In a frictionless environment, the multiplicity of equilibria is eliminated by requiring that the matching be "stable", i.e., that there is no profitable pairwise deviation. In our framework, ensuring "stability" would require that all males know the location of all females to be able to check all pairwise deviations. Since locating agents is very costly in our model, the equilibrium outcome generically does not satisfy "stability."

The result of Theorem 2 is intuitive. Recall that there are two motives for female x to target male y. The productive and the strategic motive. The payoff of a female depends on the product of the portion she appropriates from the output of the match and the probability of reciprocation. While her portion of the split does not depend on equilibrium strategies, the strategic motive does. When the search cost, θ_x, is very low, females (and males) are able to place a high probability of targeting one counterparty and exclude all others. As a result, when θ_x is extremely low, the strategic motive dominates.

It does not matter what portion of the payoff female x will get from a match with male y if the male chooses not to consider female x. When the strategic motive dominates, multiplicity of equilibria is a natural outcome. In the extreme, any pairing of agents is an equilibrium since no one has an incentive to deviate from any mutual reciprocation.

As θ_x and θ_y increase, probability distributions become less precise, as it is increasingly costly to target a particular counterparty. That is, the search costs dampen the strategic motive and the productive motive plays a bigger role. At some threshold level of θ each agent will be exactly indifferent between following the strategic motive and seeking a better match. This level of costs is precisely characterized by the "diagonal dominance" conditions of Theorem 2. Agents require the strategic motive, characterized by the off-diagonal element of the Hessian of the game, to be lower than the productive motive, captured by the diagonal element. Above the threshold the unique equilibrium has the property that each agent places a higher probability on the counterparty that promises a higher payoff, i.e., the productive motive dominates.

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

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

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

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

Equilibrium conditions (3) have an intuitive interpretation. They predict that the higher the female's private gain from matching with a male, the higher the probability of targeting that male. Similarly, the higher the probability that a male targets a particular female, the higher the probability that that female targets that male. Overall, females target males that promise higher expected private gains by placing higher probabilities on those males. Males are naturally sorted in each female's strategy by the probability of the female targeting each male. The strategies of males have the same properties due to the symmetry of the problem.

In equilibrium, a male's strategy is the best response to the strategies of females, and a female's strategy is the best response to the strategies of males. Theorem 2 predicts that an increase in θ reduces the complementarities between search strategies of females and males. Once θ is sufficiently high, the intersection of best responses leads to a unique equilibrium. Note that, by the nature of the index theorem used in the proof of uniqueness, it is enough to check diagonal dominance conditions locally in the neighborhood of the equilibrium. There is no requirement for them to hold globally. This suggests a simple way of finding equilibria of our model in most interesting cases. We first need to find one solution to the first-order conditions (3) and then check that diagonal dominance conditions are satisfied.

Now, consider the properties of equilibria for two limiting cases. First, as the search costs go to zero, targeting strategies become more and more precise. In the limit, in every equilibrium each female places a unit probability on a particular male, and that male responds with a unit probability of considering that female. Each equilibrium of this kind implements a matching of the classical assignment problem (not all of them are stable).

Second, in the opposite case when search costs go to infinity, optimal

strategies of males and females approach a uniform distribution. This unique equilibrium implements the standard uniform random matching assumption extensively used in the literature. Thus, the assignment model and the random matching model are special cases of our targeted search model, when θ is either very low or very high.

2.3 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 males would be infinitely precise.

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

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

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

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

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

$$q_y^o(x) = \exp\left(\frac{\Phi_{xy} p_x^o(y)}{\theta_y}\right) / \sum_{m'=1}^{M} \exp\left(\frac{\Phi_{x'y} p_{x'}^o(y)}{\theta_y}\right). \tag{4}$$

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

This result is reminiscent of goods with positive externalities where the producer undersupplies the good if she is not fully compensated by the marginal social benefits that an additional unit of the good would provide to society. In our model, additional search effort exerted by an individual male or female has a positive externality on the whole matching market.

For instance, when a male chooses to increase his search effort, he can better locate his preferable matches. As a consequence, the females he targets will benefit (through an increase in the personal matching rate) and the females he does not target will also be better off as his more targeted strategy will help them exclude him from their search (through a decrease in the personal matching rate). Nevertheless, in this environment agents can not appropriate all the social benefits (the output of a match) they provide to society when increasing their search effort. They only get a fraction of the payoff. This failure of the market to fully compensate both females and males

with their social marginal products leads to under-supply of search effort by both sides in the decentralized equilibrium.

Because the social gain is always the sum of private gains, there is no feasible way of splitting the payoff such that it implements the social optimum. When θ is finite and positive, a socially optimal equilibrium has to satisfy the following conditions simultaneously:

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

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

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

1) cost functions are increasing and convex;

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

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

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

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

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

Proof. See Appendix A. □

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

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

Conditions 3 and 4 together require heterogeneity to be two-sided. If heterogeneity is one-sided, i.e. condition 3 or condition 4 is violated, then the allocation of intentions towards the homogeneous side of the market will be uniform. In this case, search becomes one-sided and equilibrium allocations are efficient contingent on the actively searching side appropriating 100 percent of the payoff.[10]

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

Another way of thinking about the inefficient quantity of matches is to consider the reduction in strategic complementarities. To illustrate these complementarities consider the case of a female who chooses her strategy under the assumption that all males implement socially optimal strategies. Because a female considers only her private gains from matching with a male, the female's optimal response would be to target less accurately the best males than is socially optimal. In a second step, taking as given these strategies of females, males will be discouraged to reduce noise further, not only by the fact that they appropriate fractions of the total gains from a match, but also by the fact that females do not choose strategies that are as targeted as is socially optimal. These complementary disincentives will lower the probabilities of males pursuing their best match. Iterating in this way on

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

strategies of males and females, at each step the probability of targeting the best matches decreases. As a result, agents will target their *better* matches instead of the best possible matches.

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

2.4 Implications for Sorting

To better understand the effect of the productive and strategic motives, it is useful to consider simple examples of payoffs to understand the relative importance of these motives for equilibrium matching rates. Let us consider a matching market where there are just two males and two females, with types labeled, high (H) and low (L). Let us also consider two specific cases of the form of the payoff function.

Case one: The high type female is better off with a high type male, and the low type female is better off with a low type male. The same property is true for males. We shall generally refer to a payoff function where for each type the best option on the other side is different - as the case of *relative* advantage. Case two: Both females prefer the high type male, and both males prefer the high type female. We shall generally refer to a payoff for which everyone's best option is the same type as the case of *absolute* advantage.

In the case of relative advantage, the strategic and the productive motives

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

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

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

This last result is in stark contrast with the literature on optimal assignment, which predicts a PAM equilibrium as the only stable outcome. The

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

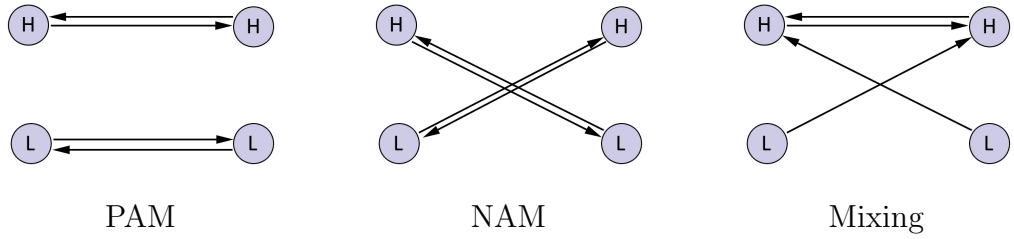

Figure 1: Three types of equilibria and their sorting patterns

Note: We show by an arrow the direction in which each agent places the highest probability.

This basic intuition has important implications for empirical inference. If the productive and strategic motives are perfectly aligned, as they are in the relative advantage case, then the shape of the equilibrium matching pattern looks very similar to the shape of the payoff function. Indeed, agents will always place the highest probability on the types that give the highest payoff and we shall see a larger number of matches between those pairs of types. The presence of a conflict between these motives, as in the absolute advantage case, drives a wedge between the shape of the payoff and the shape of the matching rates. On the one hand, you should still see more matches between

pairs of types that are more productive. On the other hand, there is a large number of competing agents that would be able to compensate for the lower payoff by a higher probability of reciprocation. The main consequence of this result is that when the payoff function is such that the two motives are in conflict, the pattern of who marries whom may differ substantially from the pattern of who would be better off with whom.

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

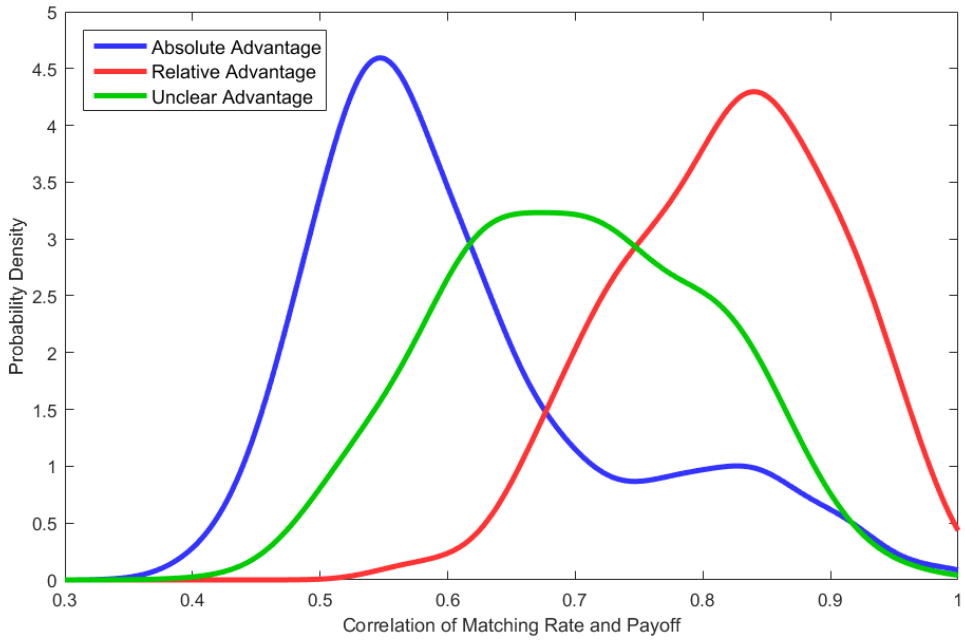

Figure 2: Correlation of Matching rate and Payoff

We find that, indeed, in the case of absolute advantage the correlation is significantly lower than that in the case of relative advantage. The intermediate shapes of payoff generate intermediate values of the correlation. Thus, when our model is the true data-generating process, the conflict between the productive and strategic motive could drive a substantial wedge between the shape of the underlying payoff function and the shape of the matching rate. Consequently, the empirical researcher could easily arrive at wrong conclusions about the shape of the underlying payoff by simply looking at the shape of the matching rates. As we shall discuss at the end of the empirical section, this is indeed what workhorse models of the marriage market do.

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

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

2.5 Invertibility

Our model builds on the interaction of strategic motives of agents and is hence more complicated computationally than leading examples in the literature. This fact has both bonuses and drawbacks. We find that in our model the mapping between the payoff and the matching rate is not invertible. By that we mean that there exist matching rate patterns that cannot in principle be generated by our model. Also, we cannot exclude the possibility that some matching rate pattern could be generated by more than one payoff function (although we could not find an example of this in practice).

This implies that our model is testable. That is, we could observe data

on matching rates that would be at odds with the predictions of our model. To illustrate this point we perform a Monte-Carlo exercise by drawing the elements of the 2-by-2 payoff matrix from a uniform distribution, computing the equilibrium and the corresponding matching rates. We normalize the total expected number of matches to one and plot all the possible vectors of equilibrium matching rates on a (3-dimensional) simplex. Figure 3 illustrates our findings.

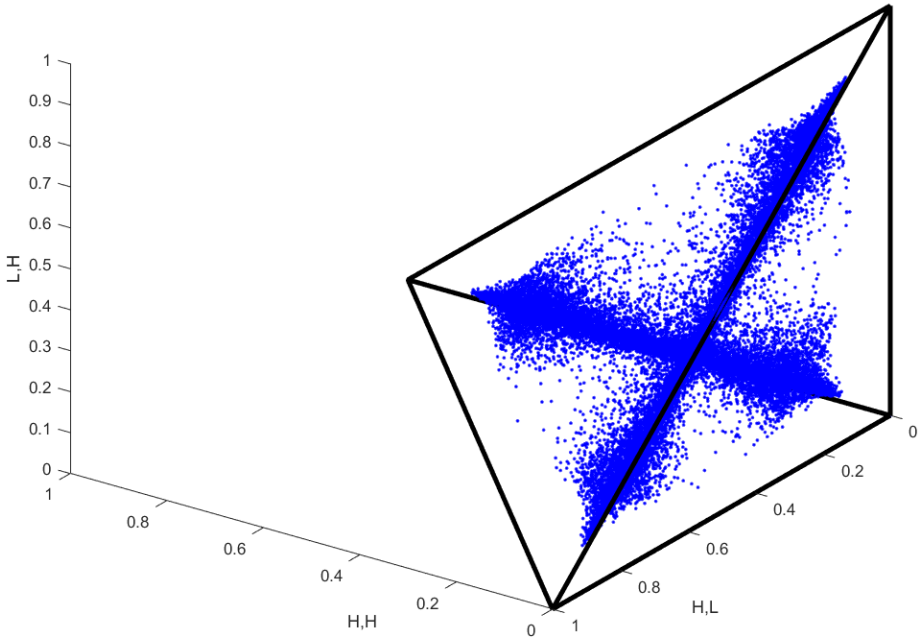

Figure 3: Subset of model-generated matching rates in a 3D simplex

We find that large white spaces remain in the simplex, implying that many shapes of matching rates cannot be obtained as an equilibrium outcome of our model. The intuition for this result is simple. If, for example, both types of males search actively, an equilibrium cannot allocate the majority of prospective matches to just one of the males and generate almost no matches for the other male. This is why we argue that our theory of targeted search is testable. It implies certain restrictions on equilibrium outcomes which may or may not be rejected by the data. This result is in contrast with workhorse models in the marriage literature, such as Choo and Siow (2006) and Galichon and Salanie (2012), which can rationalize any observed matching rates.

Given the non-invertibility of the mapping between the payoff and the matching rates, how do we test the model and estimate the payoff function from matching rate data? Our empirical methodology proceeds in three steps. First, we make a few identifying assumptions. In particular, we assume that search costs are identical across agents on both sides of the marriage market, $\theta_x = \theta_y = \theta$. This assumption will allow us to identify the ratio of the payoff to search cost, Φ_{xy}/θ, for each pair of types. In addition, we assume that each payoff is split equally between males and females, i.e. $\epsilon_{xy} = \eta_{xy} = \Phi_{xy}/2$.

Second, for any shape of the payoff function, Φ_{xy}, we find all equilibria (if there are more than one) and compute all corresponding equilibrium matching rates implied by the model. Third, we search for a shape of the payoff that maximizes the likelihood function of the data given the predicted matching rates. Whenever a proposed payoff function produces multiple equilibria, we select the one that fits the observed matching rate best, i.e. has the highest likelihood. Maximization of the likelihood function efficiently minimizes the properly weighted sum of distances between the data and the model's

prediction and should lead to consistent estimates. Maximum likelihood estimation of discrete games with multiple equilibria have been reasonably well studied in the literature, e.g. Aguirregabiria and Mira (2007). Here we do not employ any computational tricks since the 3-by-3 case can be computed by brute force in reasonable time. The results of such estimation can be treated as an upper bound on the explanatory power of the model. In the empirical section, we apply this method to three prominent examples of sorting in the marriage market and find that the model fits the data very well.

3 Empirical Application

To take the model to the data, we use a standard dataset for matching rates in the U.S. marriage market. The data on unmarried males and females and newly married couples comes from IPUMS for the year 2001.[11] For computational transparency we attribute both males and females to three equally sized bins, which we refer to as low (L), medium (M), and high (H) types. We consider three dimensions along which males and females evaluate each other in the marriage market: income, age, and education. In each case we choose the cutoffs between bins in such a way as to split the whole U.S. population of each gender to equally sized bins.

In the case of age, we restrict our attention only to adults between the ages 21 and 33. To make them as close as possible to equal size, the bins correspond to ages 21-23, 24-27, 28-33. We discard all younger and older people from the analysis because there is a disproportionate amount of unmarried people in these other age categories who only rarely marry. One reason for this may be that a large fraction of them are not searching for a spouse and are thus not participating in the marriage market. To avoid misspecification

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

due to our inability to observe search effort, we exclude them from our analysis. In the case of education, the natural breakdown into three bins is to have people who never attended college, those who are currently in college, and those who have graduated from college. Income is a continuous characteristic, so the three bins correspond to people with low, medium, and high incomes.

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

The matching rate for the case of income is presented in the left panel of Figure 4. The estimate of the underlying payoff is shown in the right panel of the same Figure. A notable property of the payoff is that it exhibits strong absolute advantage. That is, marrying a spouse with a higher income is always better. We find that the matching rate and the payoff have a correlation of 0.72.

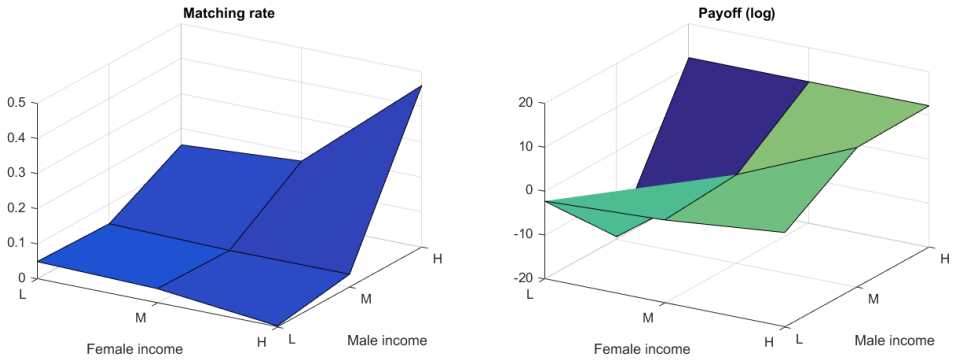

Figure 4: Sorting by income

The matching rate for the sorting by age is presented in the left panel of Figure 5. Looking at the shape of the matching rate, we would expect to see the pattern of relative advantage here, with slightly older males looking for slightly older females. However, the shape of the payoff that best explains this sorting pattern is very close to absolute advantage. Females have a strong preference for older males independent of their own age. Meanwhile, males are virtually indifferent to the age of their spouse. The highest payoff is produced by males at age 30 marrying females at age 23. The correlation between the matching rate and the payoff is a staggeringly low 0.42.

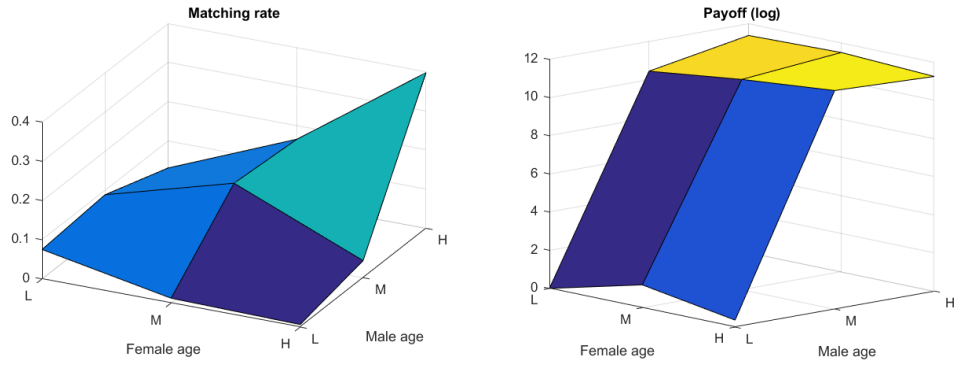

Figure 5: Sorting by age

The matching rate for sorting by education is presented in the left panel of Figure 6. In this case the payoff exhibits a combination of absolute and relative advantage. Low educated people and high educated people prefer someone with their same level of education, generating a region of relative advantage. However, people with a medium level of education tend to prefer highly educated people, generating a region with absolute advantage. The matching rate and the payoff function have a correlation of 0.52.

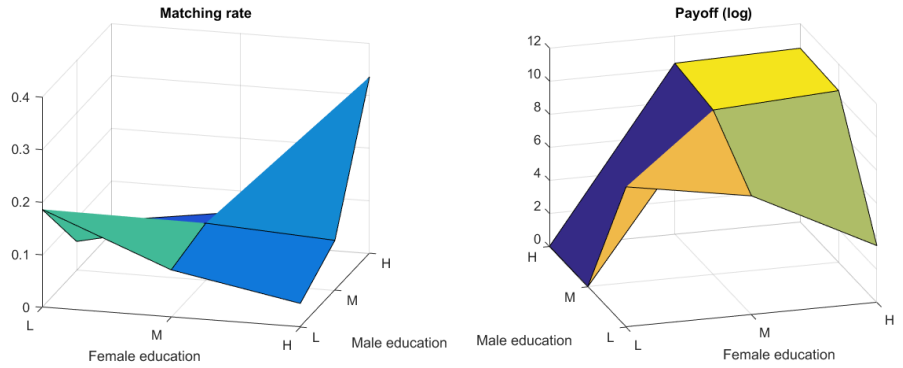

Figure 6: Sorting by education

A widely used workhorse model in the marriage literature is the model of Choo and Siow (2006). They estimate a static transferable utility model that

generates a nonparametric marriage matching function. This model postulates that, in equilibrium, each pair of cohorts of males and females reaches an implicit agreement on the matching rate among themselves; matching (or staying single) is a voluntary decision. In their model, the payoff is recovered as a simple algebraic function of the matching rates and the number of people searching. The first notable property of this mapping is that it is one-to-one, i.e., for any payoff there is a unique matching rate and for any matching rate one can invert the relationship to compute the payoff.

The second notable property is that the matching rate depends only on the characteristics of the agents directly involved in the match, but not on the characteristics of other agents present in the marriage market. This is because the strategic motive is absent from their model, so the shape of the matching rate mimics closely the shape of the payoff function. An important consequence of these two properties is that any set of matching rates observed in the data can be rationalized by some form of payoff function. Thus, the model of Choo and Siow does not place any constraints on the data and cannot be tested. This also implies that the distance between the assumptions and implications is minimal: the correlation between the matching rates across pairs of types and the implied values of the payoff are close to one.

We illustrate this feature in Figure 7 where we use the 3-by-3 Monte Carlo simulation from Section 4.1. We plot the correlation between the true underlying payoff and the equilibrium matching rate obtained from our model on the horizontal axis and the correlation between the same matching rate and the corresponding payoff function recovered by the model of Choo and Siow on the vertical axis. We find that in many cases, the shape of the true payoff and of the matching rate descends to 0.4, while the model of Choo and

Siow would imply that they have a similar shape with a correlation above 0.75. We color the payoffs with the three patterns of advantage in three different colors. We find that while the correlation depends significantly on the pattern of advantage in our model, in Choo and Siow's model it does not.

The Figure also compares our empirical findings with the Monte Carlo simulation. We find that the three prominent empirical examples that we have considered indeed belong to the range of correlation values commonly generated by payoffs that exhibit absolute advantage.

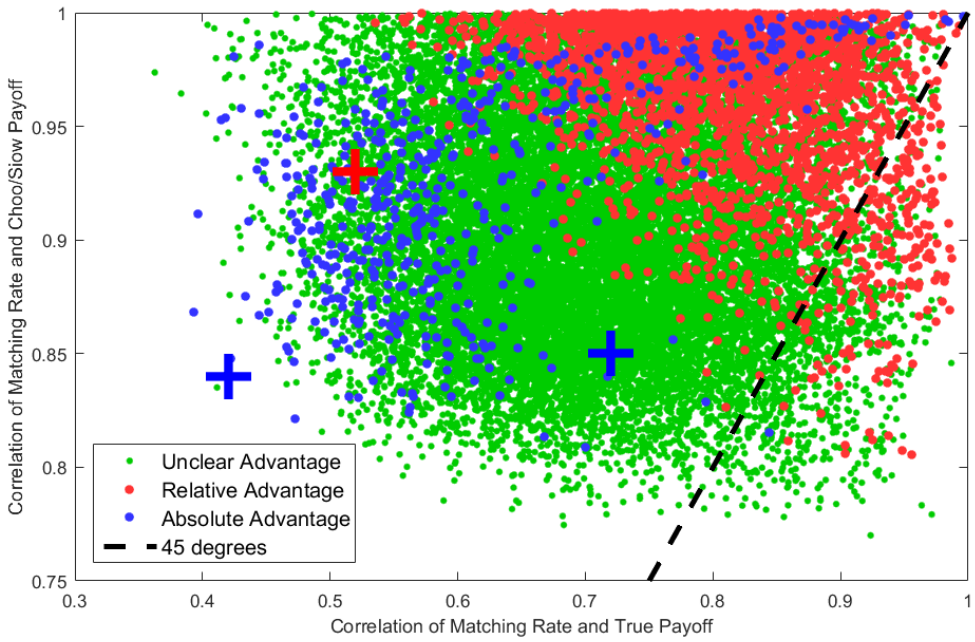

Figure 7: Monte Carlo results and Data

This result emphasizes the importance of considering the effect of strategic motives on the sorting patterns in empirical research. If a researcher looks at the data through the lens of a model with exogenous randomness, that model by construction ignores any strategic considerations that may affect agents' search strategies. As we have shown, strategic considerations can drive a significant wedge between the shape of the productive complementarities and the shape of the observed sorting pattern. Ignoring endogenous randomness may thus lead to vastly misleading conclusions regarding the amount of mismatch present in a market and the size of the losses associated with it.

4 Final Remarks

In this paper we endogenize the degree of randomness in the matching process by proposing a model where agents must pay a cost for searching to better locate potential matches. If they increase search effort, they increase the probability of targeting a better match. The model features a productive motive that drives agents to target the person that renders a higher payoff and a strategic motive that drives agents to target the person with whom their interest is more likely to be reciprocated. We believe that ignoring these considerations may result in misleading implications about the degree of mismatch present in the market and hence about the size of the losses associated with it.

With endogenous randomness as the driving force of matching patterns, our model is well suited to study a host of real-life matching markets where people typically have limited time and ability to reduce noise. Roth and Sotomayor (1990) and Sönmez and Ünver (2010) provide examples of such markets. Moreover, for many markets, equilibrium outcomes are neither pure

random matching nor optimal assignment, as documented in the empirical literature. Our model can be a useful tool for analyzing these markets.

Furthermore, our model describes markets where the degree of centralization is fairly low. This structure encompasses a number of markets ranging from labor markets to education and health care. 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 the study of the optimal design of centralization in two-sided search environments as an exciting area of future research and a practical application of our theory with far-reaching consequences.

References

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

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

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

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

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

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

[7] **Diamond, P. A., (1982).** "Wage Determination and Efficiency in Search Equilibrium." *Review of Economic Studies*, 49(2), pp. 217-227.

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

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

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

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

[12] **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.

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

[14] **Roth, A. E. and M. Sotomayor, (1990).** "Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis." *Econometric Society Monograph Series*, Cambridge University Press, 1990.

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

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

[17] **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.

[18] **Sönmez, T. and M. U. Ünver, (2010).** "Matching, Allocation, and Exchange of Discrete Resources," in J. Benhabib, A. Bisin, and M. Jackson (eds.), *Handbook of Social Economics*, Elsevier.

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

[20] **Vives, X. (1990).** "Nash Equilibrium with Strategic Complementarities." *Journal of Mathematical Economics*, 19, pp. 305-321.

NOT FOR PUBLICATION

Appendix A: Proof of Theorem 3

The proof proceeds in 3 steps.

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

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

$$\text{POFOC}_{q_y(x)}: \qquad \Phi_{xy}\tilde{p}_x(y) - \left.\frac{\partial c_y(\tilde{\kappa}_y)}{\partial \tilde{\kappa}_y}\right|_{\tilde{q}_y(x)} \frac{1}{\ln 2}\left(\ln \frac{\tilde{q}_y(x)}{1/F} + 1\right) - \tilde{\lambda}_y = 0$$

$$\text{POFOC}_{p_x(y)}: \qquad \Phi_{xy}\tilde{q}_y(x) - \left.\frac{\partial c_x(\tilde{\kappa}_x)}{\partial \tilde{\kappa}_x}\right|_{\tilde{p}_x(y)} \frac{1}{\ln 2}\left(\ln \frac{\tilde{p}_x(y)}{1/M} + 1\right) - \tilde{\lambda}_x = 0$$

$$\text{CEFOC}_{q_y(x)}: \qquad \eta_{xy}p_x(y) - \left.\frac{\partial c_y(\kappa_y)}{\partial \kappa_y}\right|_{q_y(x)} \frac{1}{\ln 2}\left(\ln \frac{q_y(x)}{1/F} + 1\right) - \lambda_y = 0$$

$$\text{CEFOC}_{p_x(y)}: \qquad \varepsilon_{xy}q_y(x) - \left.\frac{\partial c_x(\kappa_x)}{\partial \kappa_x}\right|_{p_x(y)} \frac{1}{\ln 2}\left(\ln \frac{p_x(y)}{1/M} + 1\right) - \lambda_x = 0$$

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

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

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

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

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

and

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

Denote them a_y and a_x respectively.

It then follows that:

$$\begin{aligned}
\Phi_{xy}\tilde{p}_x\left(y\right) - \tilde{\lambda}_y &= \left.\frac{\partial c_y\left(\tilde{\kappa}_y\right)}{\partial \tilde{\kappa}_y}\right|_{\tilde{q}_y(x)} \frac{1}{\ln 2}\left(\ln \frac{\tilde{q}_y\left(x\right)}{1/M} + 1\right) \\
&= \left.\frac{\partial c_y\left(\kappa_y\right)}{\partial \kappa_y}\right|_{q_y(x)} \frac{1}{\ln 2}\left(\ln \frac{q_y\left(x\right)}{1/M} + 1\right) \\
&= \eta_{xy}p_x\left(y\right) - \lambda_y
\end{aligned}$$

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

$$(i) \qquad M\exp\left(1 + \tfrac{\tilde{\lambda}_y}{a_y/\ln 2}\right) = \sum_{x=1}^{M}\exp\left(\tfrac{\Phi_{xy}p_x(y)}{a_y/\ln 2}\right)$$

$$(ii) \qquad M\exp\left(1 + \tfrac{\lambda_y}{a_y/\ln 2}\right) = \sum_{x=1}^{M}\exp\left(\tfrac{\varepsilon_{xy}(x)p_x(y)}{a_y/\ln 2}\right)$$

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

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

$$\frac{\sum_{x'=1}^{M}\exp\left(\frac{\Phi_{x'y}p_{x'}(y)}{a_f/\ln 2}\right)}{\sum_{m'=1}^{M}\exp\left(\frac{\varepsilon_{x'y}p_{x'}(y)}{a_y/\ln 2}\right)} = \frac{\exp\left(\frac{\Phi_{xy}p_x(y)}{a_y/\ln 2}\right)}{\exp\left(\frac{\varepsilon_{xy}p_x(y)}{a_y/\ln 2}\right)} \qquad \text{for all } x$$

Hence,

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

Therefore, either:

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

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

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

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

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

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

Appendix B: One-sided model

Here we consider a one-sided model where females are searching for males who are heterogeneous in type and females face a search cost. We assume that there is no heterogeneity on the female side of the market. As such the probability that a male reciprocates the intentions of a female is given by q_y. The strategy of a female, denoted $p_x(y)$, represents the probability of female x locating male y. It is also the female's probability distribution. We assume that each female can rationally choose her strategy facing a trade-off between a higher payoff and a higher cost of searching.

Like before a female's cost of searching is given by $c_x(\kappa_x)$. Once the optimal distribution $p_x(y)$ is chosen, each female draws from it to determine which male to contact.[12]

Female x chooses a strategy $p_x(y)$ to maximize her expected income flow:

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

We normalize the outside option of females to zero. A female receives her expected share of the payoff in a match with male y conditional on matching with that male. She also incurs a cost that depends on search effort:

$$\kappa_x = \sum_{y=1}^{M} p_x(y) \ln \frac{p_x(y)}{1/M} \tag{5}$$

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

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

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

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

[12] As in the model of Section 2, we assume that each female pursues only one male.

When in addition the cost functions are differentiable, it is easy to verify that first-order conditions are necessary and sufficient conditions for equilibrium.[13] Rearranging the first order conditions for the female, we obtain:

$$p_x^*(y) = \exp\left(\frac{\varepsilon_{xy} q_y}{\frac{1}{\ln 2}\left.\frac{\partial c_x(\kappa_x)}{\partial \kappa_x}\right|_{p_x^*}}\right) / \sum_{y'=1}^{M} \exp\left(\frac{\varepsilon_{xy'} q_{y'}}{\frac{1}{\ln 2}\left.\frac{\partial c_x(\kappa_x)}{\partial \kappa_x}\right|_{p_x^*}}\right). \tag{6}$$

This is an implicit relationship as p_x^* appears on both sides of the expression. If cost functions are linear functions of , κ_x, then the derivatives on the right hand side are independent of p_x^*, and the relationship becomes explicit.

The equilibrium condition (6) has an intuitive interpretation. It predicts that the higher is the female's expected gain from matching with a male, the higher is the probability placed on locating that male. Thus, males are naturally sorted in each female's strategy by probabilities of contacting those males.

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

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

[13] Taking derivatives of the Lagrangian function corresponding to the problem of female x, we obtain for all y:

$$\varepsilon_{xy} q_y - \left.\frac{\partial c_x(\kappa_x)}{\partial \kappa_x}\right|_{p_x^*} \frac{1}{\ln 2}\left(\ln \frac{p_x^*(y)}{1/M} + 1\right) = \lambda_x$$

We can invert this first-order condition to characterize the optimal strategy:

$$p_x^*(y) = \frac{1}{M} \exp\left(\frac{\varepsilon_{xy} q_y - \lambda_x}{\frac{1}{\ln 2}\left.\frac{\partial c_x(\kappa_x)}{\partial \kappa_x}\right|_{p_x^*}} - 1\right).$$

subject to (5) and to the constraint that the $p_x(y)$'s are well-defined probability distributions. Under the assumption of increasing convex cost functions, the social welfare function is concave in the strategies of females. Hence, first-order conditions are sufficient conditions for a maximum. Rearranging and substituting out Lagrange multipliers, we arrive at the following characterization of the social planner's allocation:

$$p_x^o(y) = \exp\left(\frac{\Phi_{xy} q_y}{\frac{1}{\ln 2} \left.\frac{\partial c_x(\kappa_x)}{\partial \kappa_x}\right|_{p_x^o}}\right) / \sum_{y'=1}^{M} \exp\left(\frac{\Phi_{xy'} q_y}{\frac{1}{\ln 2} \left.\frac{\partial c_x(\kappa_x)}{\partial \kappa_x}\right|_{p_x^o}}\right). \tag{7}$$

The first observation to make is that the structure of the social planner's solution is very similar to the structure of the decentralized equilibrium. Second, from the female's perspective, the only difference between the centralized and decentralized equilibrium strategies is that the probability of locating a male depends on the social gain from a match rather than on the private gain. Thus, it is socially optimal to consider the whole expected payoff when determining the socially optimal strategies, while in the decentralized equilibrium females only consider their private gains.

To decentralize the socially optimal outcome the planner needs to give all of the payoff to the females, $\varepsilon_{xy} = \Phi_{xy}$, effectively assigning them a share of 1. Note that, if the planner could choose the probability that a male reciprocates a female, q_y, he would also set it to 1.

The only special cases, when the outcome is always efficient are the limiting cases discussed earlier. When search costs are absent, the equilibrium of the model is socially optimal. When costs are very high, the random matching outcome is the best possible outcome. For all intermediate values of costs, the decentralized equilibrium is socially efficient contingent on the female having all the bargaining power.