International Journal of Game Theory manuscript No.
(will be inserted by the editor)

Instability of matchings in decentralized markets with various preference structures

Kimmo Eriksson[1], **Olle Häggström**[2]

[1] Mälardalen University, e-mail: kimmo.eriksson@mdh.se

[2] Chalmers University of Technology

The date of receipt and acceptance will be inserted by the editor

Correspondence to: Kimmo Eriksson, Mälardalen University, IMA, Box 883, SE-721 23 Västerås, Sweden

Abstract In any two-sided matching market, a stable matching can be found by a central agency using the deferred acceptance procedure of Gale and Shapley. But if the market is decentralized and information is incomplete then stability of the ensuing matching is not to be expected. Despite the prevalence of such matching situations, and the importance of stability, little theory exists concerning instability. We discuss various measures of instability and analyze how they interact with the structure of the underlying preferences. Our main result is that even the outcome of decentralized matching with incomplete information can be expected to be "almost stable" under reasonable assumptions.

Key words stable matching – blocking pair – instability – preference structure – decentralized market – maximin matching

1 Introduction

A matching is *stable* if there is no *blocking pair*, that is, a pair where both agents would prefer each other to their partners in the matching. Thus, whether a given matching is stable is determined by the preferences of agents. Does every preference structure allow a stable matching? It has long been known that the answer to this question hinges on whether the market is two-sided. Indeed, the seminal result in the theory of two-sided matching is that the deferred acceptance procedure of Gale and Shapley (1962) always yields a stable matching, regardless of the preferences held by the agents (cf. Roth and Sotomayor, 1990).

A matching procedure like the deferred acceptance procedure (DAP) is, of course, easily implemented in any centralized market – but its applicability is even greater. As argued by Roth and Xing (1997), the DAP can be followed also by agents in a decentralized market if the two sides have distinct roles and every agent has complete knowledge of his or her own preferences. However, it is often the case that agents do *not* know their own preferences from the beginning. This is the problem of *mate search* in biology and social psychology, be it search for a sexual mate or a person to chat with at a cocktail party. In such markets, unless search is exhaustive it is impossible to guarantee stability of any matching obtained. Still, by making greater search efforts we would expect agents to find a matching that is somehow closer to stable. This is the general background to our two overarching research questions: (1) How can instability of matchings be measured so that it becomes meaningful to speak of one matching being "closer to stable" than another? (2) In decentralized matching markets with incomplete information, is it true that we can expect agents to match up in a way that is "close to stable"?

The existing game theoretic literature bearing upon our question is quite small. We will briefly discuss three studies we have identified as most relevant.

Roth and Xing (1997) simulated the American entry-level market for clinical psychologists with agents following the rules of APPIC, the organization that administers the market. Agents in this market have full knowledge of their own preferences. Roth and Xing measured the degree of instability of the final matchings by the proportion of blocking agents, that is, agents that belong to some blocking pair. They found that the degree of instability depended on the preference structure (common or random).

Ünver (2005) compared a decentralized market with various centralized market mechanisms in the laboratory, following an earlier experiment by Kagel and Roth (2000). The market consisted of three "high types" and three "low types" at each side, with everyone preferring a high type to a low type. Thus preferences were common, and every stable outcome has high types matched to high types and low types to low types, so a high type matched to a low type was counted as a "mismatch." Ünver measured instability as the proportion of unmatched and mismatched agents, and found for the decentralized market a total of 40 percent of agents in these categories. This instability was significantly higher than when centralized market mechanisms were used.

Niederle and Roth (2006) conducted an experiment where "applicants" tried to match up with "firms" in a procedure during which qualities were gradually revealed. Various treatments were compared, differing in whether offers were "open" (can be put on hold by applicant) or "exploding" (must be immediately accepted or rejected). Among other things, the functioning of the market was measured by the number of blocking pairs. The authors found that the number of blocking pairs was larger when exploding offers were allowed.

To summarize this brief review of the literature, instability has been measured in various ways by counting either blocking agents, or "mismatched" agents, or blocking pairs. Instability of matchings has been found to vary with several aspects of the underlying market, including the preference structure and the options available to agents.

In this paper we will argue that the proportion of blocking pairs among all possible pairs is usually the best measure of instability. The exception is when we want to be able to compare instability of matchings across different preference structures, in which case we will show that the blocking pairs measure is biased. We will also suggest a way to eliminate this bias.

We will then discuss what degree of instability can be expected in the outcome of a decentralized matching market with incomplete information. If agents in effect just reach a random matching, we show that the expected instability is substantial. However, if each agent acts according to a quite reasonable and simple heuristic, then we show that our measure of instability tends to zero as the number of agents grow. Thus, it seems that decentralized markets would indeed be able to yield matchings that are close to stable.

Outline of this paper

In Section 2 we define four special preference structures of interest: *random*, *common*, *homotypic* and *antithetical* preferences. In the next section we discuss several possible measures of instability, including a novel measure designed to avoid bias in comparisons across preference structures. In Section 4 we investigate how these instability measures interact with preference structures. Homotypic resp. antithetical preferences turn out to be the extreme cases, maximizing resp. minimizing the expected number of blocking pairs in a random matching. Finally, in Section 5, we give a mathematical proof that if agents use a simple heuristic in a decentralized mate search situation, the instability of the resulting matching tends to zero with increasing size of the market.

1.1 Acknowledgments

This work was partially supported by a grant to the first author from the Swedish Research Council.

2 Markets and preference structures

We will always assume markets to be two-sided, each side (X and Y, or women and men) consisting of n agents so that a complete matching is possible. However, a similar theory can be developed for one-sided markets, so called "roommate problems" (Eriksson and Strimling, unpublished).

A *preference structure* will be an n-by-n matrix $\mathbf{P}^{(n)}$ where each entry $\mathbf{P}^{(n)}_{xy}$ is a pair (A_{xy}, B_{xy}). The first number signifies how agent x ranks y among all Y-agents (where 0 is the worst possible rank and $n-1$ is the best possible rank). Similarly, the second number signifies how agent y ranks x among all X-agents. Hence, in every row (resp. column) the first (resp. second) numbers of the pairs constitute a permutation of the integers from 0 to $n-1$.

In models of mate search one must make an assumption about the underlying preference structure. There seem to be three standard options. Most common is an assumption of *common* preferences, i.e. a universally held notion of attractiveness (cf. Todd and Miller, 1999; Alpern and Reyniers, 2005). When common preferences are contested, the usual alternative put forward is *homotypic* preferences, where everyone prefers someone like themselves (cf. Alpern and Reyniers, 1999). A third alternative is to model agents' preferences as *random*, i.e. independent of each other (cf. Eriksson, Sjöstrand and Strimling, 2006). Sometimes these assumptions are pitted against each other; Roth and Xing (1997) compared decentralized matching in common and random preference structures, whereas the question of the classic paper by Kalick and Hamilton (1986) can be formulated as whether real human mating preferences are common or homotypic.

In this paper we will add a fourth fundamental preference structure that will prove interesting. Say that preferences are *antithetical* if a man likes a woman more the less she likes him (and vice versa). As Groucho Marx quipped: "I don't want to belong to any club that will accept me as a member." Although not a conventional assumption in the scientific literature, versions of this model are common in popular culture and it does not seem to be alien to human psychology.

We will now give exact definitions of these preference structures.

Definition 1 *Preferences are* common *if agents can be numbered so that* $A_{xy} = y$ *and* $B_{xy} = x$ *for all* $x \in X$ *and* $y \in Y$.

Preferences are homotypic *if* $A_{xy} = B_{xy}$ *for all* $x \in X$ *and* $y \in Y$.

Preferences are antithetical *if* $A_{xy} = n - 1 - B_{xy}$ *for all* $x \in X$ *and* $y \in Y$.

Preferences are random *if for each* $x \in X$ *her preference function* $\left(A_{xy}\right)_{y=1}^{n}$ *is independently drawn at random with uniform probability from the set of all permutations (and similarly for all* Y*-agents).*

We will use the notation $\mathbf{P}_{\mathrm{com}}^{(n)}$, $\mathbf{P}_{\mathrm{hom}}^{(n)}$, $\mathbf{P}_{\mathrm{ant}}^{(n)}$ *resp.* $\mathbf{P}_{\mathrm{ran}}^{(n)}$ *for a market of size* n *where preferences are common, homotypic, antithetical, resp. random.*

2.1 Preference structures and latin squares

Both homotypic and antithetical preference structures are in bijection with latin squares, since the conditions boil down to there being exactly one A-value of each sort in every row and column, with B-values completely determined by the A-values. It is well-known that the number of nonisomorphic latin squares of size n quickly becomes very large. On the other hand, up to isomorphism, there is just one instance with common preferences for a given value of n.

2.2 Uniqueness of stable matchings for common and homotypic preferences

Usually there are many different stable matchings in a market. Pittel (1989) showed that under random preferences the expected number of stable matchings grows as $n \ln n$. However, some special preference structures have a unique stable matching. For instance, it is well-known and trivial to show that under common preferences the only stable matching is the perfectly assortative matching where all pairs consist of equally ranked agents. Similarly, under homotypic preferences the unique stable matching has every agent matched with his or her highest-ranked partner.

The case of antithetical preferences is not as well-known and several stable matchings exist. For instance, for any value of i between 0 and $n-1$, we evidently obtain a stable matching by letting every woman have her ith ranked man (who then has his $(n-i)$th ranked woman). Sometimes there are many more stable matchings, see the book by Gusfield and Irving (1989, section 1.3.2).

3 Measuring instability

As mentioned in the introduction, many measures of instability of matchings are possible. For instance, Ünver (2005) simply counted the mismatched agents compared to the unique stable matching under the common preferences used. This measure is not generally applicable, however, since most preference structures allow many different stable matchings. In this respect, it is better to follow Roth and Xing (1997) in counting blocking agents instead of mismatched agents. But although the existence of blocking agents certainly makes the matching unstable, the number of blocking agents does not really tell us the degree of instability since we do not know the number of blocking partners of any agent; the greater this number, the more likely that this agent will at some point discover and exploit the instability.

For these reasons, we believe that the best measure of instability of matchings is obtained by counting the blocking pairs, like Niederle and Roth (2006). We will divide the number of blocking pairs by n^2 (the total number of possible pairs) to obtain the probability that two random agents will prefer each other to their partners in the matching. This seems to us a very clear notion of instability.

Definition 2 *For any matching μ under preference structure $\mathbf{P}^{(n)}$ on a set of n agents, let $B_{\mathbf{P}}^{(n)}(\mu)$ denote the number of blocking pairs. Let $\hat{B}_{\mathbf{P}}^{(n)}(\mu)$*

denote the proportion of blocking pairs:

$$\hat{B}_{\mathbf{P}}^{(n)}(\mu) = B_{\mathbf{P}}^{(n)}(\mu)/n^2.$$

We will call this number the **instability** *of the matching μ.*

There is one caveat with this measure. In an experimental study, Eriksson and Strimling (unpublished) compare the outcomes of a decentralized matching market across different preference structures. As we will show below, our instability measure is biased in the sense that the instability of a random matching can be expected to be considerably greater if preferences are homotypic than if they are common or random, and even less if preferences are antithetical. In other words, it takes more skill or effort for agents to achieve low instability if preferences are homotypic. For comparisons of the instability of outcomes when the preference structure varies, we propose another measure.

Definition 3 *The* **relative instability** *of a matching μ under preference structure $\mathbf{P}^{(n)}$ is the proportion of matchings that have less instability than μ. There are $n!$ possible complete matchings, so the relative instability of μ is given by the expression*

$$\frac{|\{\mu' : B_{\mathbf{P}}^{(n)}(\mu') < B_{\mathbf{P}}^{(n)}(\mu)\}|}{n!}.$$

It is a hard combinatorial problem to find a formula for the distribution of matchings over number of blocking pairs (cf. Abraham et al, 2006, for complexity results on related problems). For small n, relative instability can be calculated by a computer through complete enumeration of all possible matchings. For larger markets, relative instability can be estimated through Monte Carlo methods.

4 Instability of random matchings

By a "random matching" in a given market we shall mean a matching drawn at random from a uniform distribution of all possible complete matchings. This is the simplest possible model of a mechanism that can be used in decentralized matching (using no information whatsoever), thus providing a baseline. We shall compute the expected instability for a random matching in each of our four fundamental preference structures, yielding the following main result.

Theorem 1 *Let μ be a random matching in a two-sided market of size n. Then*

$$\lim_{n\to\infty} E[\hat{B}_{\mathbf{P}^{(n)}_{\mathrm{hom}}}(\mu)] = 1/3,$$

$$\lim_{n\to\infty} E[\hat{B}_{\mathbf{P}^{(n)}_{\mathrm{com}}}(\mu)] = \lim_{n\to\infty} E[\hat{B}_{\mathbf{P}^{(n)}_{\mathrm{ran}}}(\mu)] = 1/4,$$

and

$$\lim_{n\to\infty} E[\hat{B}_{\mathbf{P}^{(n)}_{\mathrm{ant}}}(\mu)] = 1/6.$$

Furthermore, the homotypic resp. antithetical preference structures attain the maximal resp. minimal expected instability among all possible preference structures.

For random preferences the proof is almost trivial: The expected instability of a random matching tends to 1/4 since the probability that two random agents will prefer each other to their partners in the matching tends to 1/2 times 1/2. We will now proceed with a more general analysis.

Lemma 1 *For a given preference structure $\mathbf{P}^{(n)}$ of size n, the expected instability of a random matching μ is*

$$E[\hat{B}_{\mathbf{P}^{(n)}}(\mu)] = \sum_{x\in X, y\in Y} \frac{(n-1-A_{xy})(n-1-B_{xy})}{n^3(n-1)}.$$

Proof Consider two agents x and y. If they are already matched, then they are not a blocking pair. The probability of this happening in a random matching is $1/n$, so with probability $(n-1)/n$ the random matching does not match x with y. In this case, there are $(n-1)^2$ possible partner combinations for x and y in a matching, all equally likely. The numerator in the lemma is the number of these subcases in which x and y would prefer each other instead of the matches. By dividing the numerator with $(n-1)^2$, the total number of subcases, and multiplying the result with $(n-1)/n$, the probability of being in this case, we obtain the probability for x and y being a blocking pair:

$$\frac{(n-1-A_{xy})(n-1-B_{xy})}{n(n-1)}.$$

The expected instability is the average of these probabilities for all pairs x, y, and there are n^2 such pairs. The lemma follows.

In order to proceed we need a new version of the following classical elementary inequality: $a_1b_2 + a_2b_1 \leq a_1b_1 + a_2b_2$ if $a_1 \leq a_2$ and $b_1 \leq b_2$.

Lemma 2 *Let $\bar{a} = a_1, a_2, \ldots, a_{2m}$ be a weakly increasing number sequence of even length. For any complete matching ν among the indices $1, 2, \ldots, 2m$, let $S_\nu(\bar{a})$ be the sum of all products of matched number pairs in $\bar{a}$:*

$$S_\nu(\bar{a}) = \sum_{\{i,j\} \in \nu} a_i a_j.$$

Then $S_\nu(\bar{a})$ is maximized by the matching $\nu_{\max} = (\{1,2\}, \{3,4\}, \ldots, \{2m-1, 2m\})$ and minimized by the matching $\nu_{\min} = (\{1, 2m\}, \{2, 2m-1\}, \ldots, \{m, m+1\})$.

Proof For any quadruple $i < j < k < l$ matched in pairs, there are three possible matchings depending on whether i goes with j, k or l. From the classical inequality it follows that

$$a_i a_l + a_j a_k \leq a_i a_k + a_j a_l \leq a_i a_j + a_k a_l.$$

Starting in any matching we can stepwise increase (resp. decrease) the product sum by rematching every matched quadruple $i < j < k < l$ so that i goes with j (resp. l). The global maximum (resp. minimum) of $S_\nu(\bar{a})$ is therefore attained by the unique (up to automorphisms of the weakly increasing sequence) matching $\nu_{\max}$ (resp. $\nu_{\min}$) for which all quadruples are locally maximal (resp. minimal).

We can now prove the remaining parts of Theorem 1.

Proof For any preference structure $\mathbf{P}^{(n)}$, and with (x, y) ranging over all ordered pairs of agents, the multiset $\{n - 1 - A_{xy}\}_{x \in X, y \in Y} \cup \{n - 1 - B_{xy}\}_{x \in X, y \in Y}$ consists of $2n$ copies each of the values $0, 1, \ldots, n-1$. Taking this ordered multiset as the sequence $\bar{a}$ in Lemma 2, we conclude that the sum

$$\sum_{(x,y)} (n-1-A_{xy})(n-1-B_{xy}), \tag{1}$$

over ordered pairs (x, y), is maximized by homotypic preferences and minimized by antithetical preferences.

Computing the sum (1) for each of the three preference structures in the theorem is a matter of elementary analysis, and we omit the details. For homotypic preferences we obtain:

$$n\left(0^2 + 1^2 + \cdots + (n-1)^2\right) = n^4/3 + O(n^3).$$

For antithetical preferences:

$$n\,(0 \cdot (n-1) + 1 \cdot (n-2) + \cdots + (n-1) \cdot 0) = n^4/6 + O(n^3).$$

For common preferences:

$$(0 + 1 + \cdots + (n-1))^2 = n^4/4 + O(n^3).$$

In the formula in Lemma 1, these sums are divided by $n^3(n-1) = n^4 + O(n^3)$. This gives us the respective limiting expected instabilities of 1/3, 1/4 and 1/6 for the homotypic, common and antithetical preference structures.

5 Instability when agents use a simple heuristic

The literature on human mate search commonly assume that agents use a simple heuristic to decide whether to accept a partner as a permanent mate (Todd and Miller, 1999; Simão and Todd, 2002, 2003). The basic assumptions in these models are as follows. Agents, starting out with no information, meet potential partners at random and evaluate them. Mating occurs if both parties find the match above a certain preference threshold. Thresholds are lowered with time, as agents realize they cannot be too choosy. Agents can return to previously evaluated partners as long as they have not mated already.

For simplicity we will assume that all agents follow the same threshold heuristic. If the starting threshold is sufficiently high and the market is sufficiently small, agents will gain complete information before any mating occurs. Then, as the threshold T is lowered, the first pair x, y will mate and leave the market when $A_{xy} > T$ and $B_{xy} > T$. Consequently agents participating in decentralized matching can in principle implement a matching mechanism that we may call *maximin* matching.

Definition 4 *The* maximin *matching mechanism determines a complete matching as follows. Let x, y be the pair of agents maximizing $\min\{A_{xy}, B_{xy}\}$. If there are several such pairs, choose the pair that maximizes $\max\{A_{xy}, B_{xy}\}$. If there are several such pairs, use any tie-breaking rule. Match x with y. Repeat the procedure with the remaining agents until none is left.*

We will now see that maximin matching fares surprisingly well with respect to instability of the outcome. Indeed, maximin matching finds the unique stable outcome for all the non-random fundamental preference structures.

Theorem 2 *The maximin matching mechanism will result in a stable outcome under both common, homotypic, and antithetical preferences.*

Proof We shall see that in all three cases, we can predict the outcome of the mechanism as one of the stable matchings described in Section 2.2. Under common preferences, the maximin mechanism will first match the two best agents, then the second-best agents, etc. Under homotypic preferences, it will match every agent with his or her favorite. Under antithetical preferences for odd values of n, the mechanism will match every agent with his or her average partner. For even n the outcome will depend on the tie-breaking rule; e.g. if priority is given to female preferences, then females will obtain mates of rank $n/2$ while males obtain mates of rank $n/2 - 1$.

For other preference structures, the outcome of the maximin mechanism is not necessarily stable. It is easy to construct a family of preference structures where the outcomes have instability tending to 3/8 (omitted here), and possibly even larger instabilities can occur. Still, this is the exception and not the norm. Our main result below says that with random preferences, the expected instability tends to zero.

Theorem 3 *Let μ be the outcome of maximin matching in a random preference structure $\mathbf{P}^{(n)}_{\text{ran}}$. Then*

$$\lim_{n\to\infty} E[\hat{B}_{\mathbf{P}^{(n)}_{\text{ran}}}(\mu)] = 0.$$

The remainder of this section is devoted first to proving this result, then a generalization to incomplete information.

5.1 Proof of Theorem 3

In order to analyze maximin matching we introduce some terminology and notation.

Definition 5 *The* min-rank *of a pair x, y is the value $m_{xy} = \min(A_{xy}, B_{xy})$.*

Our basic strategy to prove Theorem 3 consists of two steps: We will show that, on the one hand, an agent in a matched pair with high min-rank cannot be part of many blocking pairs; on the other hand, maximin matched pairs have almost always high min-rank. Hence there cannot be too many blocking pairs. The first step is taken by the following lemma.

Lemma 3 *Given a market of size n and a matching μ, an agent x cannot be a member of more than $n - 1 - m_{x\mu(x)}$ blocking pairs.*

Proof That (x, y) is a blocking pair to μ implies that $A_{xy} > A_{x\mu(x)} \geq m_{x\mu(x)}$. For a fixed x there can only be $n - 1 - m_{x\mu(x)}$ agents y that x rank so highly.

The second step is more difficult. Recall that by definition the maximal rank is $n - 1$. Among the matched pairs we will show that almost all min-ranks are close to maximal.

Lemma 4 *Let μ be the outcome of maximin matching in a random preference structure $\mathbf{P}^{(n)}_{\mathrm{ran}}$. Then, for arbitrarily small $\epsilon, \delta > 0$, the probability that at least ϵn of the matched pairs have a min-rank that does not exceed $(1 - \delta)(n - 1)$ tends to 0 as $n \to \infty$.*

Proof The event in the lemma implies that the $n \times n$-matrix $\mathbf{P}_{\mathrm{ran}}^{(n)}$ contains an $\epsilon n \times \epsilon n$ submatrix M such that all its entries (A_{xy}, B_{xy}) satisfy $\min(A_{xy}, B_{xy}) \leq (1-\delta)(n-1)$. Let us denote this event by D_n. Thus, it is sufficient that we prove $P(D_n) \to 0$ as $n \to \infty$.

To begin with, by simple combinatorics and Stirling's formula we can bound the number of submatrices of size $\epsilon n \times \epsilon n$ by

$$\binom{n}{\epsilon n}^2 \leq C(2\pi\epsilon(1-\epsilon)n)^{-1}(1-\epsilon)^{-2(1-\epsilon)n}\epsilon^{-2\epsilon n}, \tag{2}$$

for any $C > 1$ and sufficiently large n.

Now we wish to estimate the probability that in a given $\epsilon n \times \epsilon n$ submatrix M all $\min(A_{xy}, B_{xy}) \leq (1-\delta)(n-1)$. A fundamental fact is that, for any single entry, $P(A_{xy} > (1-\delta)(n-1)) = \delta$ and similarly for B_{xy}. However, different entries are not independent; if we examine the $\epsilon^2 n^2$ elements of M, one at a time, the conditional probability that $A_{xy} > (1-\delta)(n-1)$ depends on the values we have already observed in the same row. However, assuming (without loss of generality) that $\epsilon < \delta/2$, this conditional probability will never be less than $\delta/2$ since there will always be at least $\delta n/2$ $A_{x\cdot}$-values exceeding $(1-\delta)(n-1)$ remaining among the entries in this row that have not been examined. In combination with the same argument for the B_{xy}-value, we conclude that the conditional probability that $\min(A_{xy}, B_{xy}) \leq (1-\delta)(n-1)$ is at most $1-(\delta/2)^2$. Hence, the probability that the submatrix M has all min-ranks $\leq (1-\delta)(n-1)$ is at most $(1-(\delta/2)^2)^{\epsilon^2 n^2}$.

Finally, multiplication with the total number of submatrices from (2) yields

$$P(D_n) \leq C(2\pi\epsilon(1-\epsilon)n)^{-1}(1-\epsilon)^{-2(1-\epsilon)n}\epsilon^{-2\epsilon n}(1-(\delta/2)^2)^{\epsilon^2 n^2} \to 0$$

as $n \to \infty$.

According to our strategy we are now ready to prove Theorem 3.

Proof Let μ be the outcome of maximin matching in a random preference structure $\mathbf{P}^{(n)}_{\mathrm{ran}}$. By definition of instability, $\hat{B}_{\mathbf{P}^{(n)}_{\mathrm{ran}}}(\mu) \geq 0$. It remains to show that

$$\lim_{n\to\infty} E[\hat{B}_{\mathbf{P}^{(n)}_{\mathrm{ran}}}(\mu)] < \gamma$$

for any $\gamma > 0$. Choose $\delta, \epsilon > 0$ so that $\delta + \epsilon < \gamma$. We will now assume that at least $(1-\epsilon)n$ of the matched pairs have min-rank greater than $(1-\delta)(n-1)$. Among those agents, it follows from Lemma 3 that any x can be part of at most $\delta(n-1)$ blocking pairs, whereas any x among the remaining agents can be part of at most n blocking pairs. Hence the total number of blocking pairs is limited by $(1-\epsilon)n\delta(n-1) + \epsilon n^2 < \gamma n^2$ so the instability is limited by γ. According to Lemma 4, the assumption we made holds with probability approaching 1 as n grows to infinity; in the complementary case, the instability is limited by 1. We conclude that the expected instability is limited by γ.

5.2 Incomplete information

Let us return to our simple heuristics for decentralized matching. If the market is large, agents will not have the time or possibility to evaluate all possible partners before any mating occurs. The easiest way to incorporate this in our model is to assume that every agent first evaluates a random subset of partners, and then the maximin mechanism starts operating on this partial information. In this generalized model, we ask how many partners agents must evaluate in order for the instability of the outcome to remain low. The answer is straightforward — it is sufficient that the expected number of evaluated partners grows with n, however slowly.

Definition 6 *The p-partial maximin matching mechanism determines a complete matching from a preference structure* $\mathbf{P}^{(n)}$ *as follows. Define an incomplete preference structure* $\mathbf{P}^{*(n)}$ *by randomly "evaluating" a subset of all possible pairs such that any pair has an independent probability* p *of being "evaluated", and for each such pair* (x, y) *setting* $(A^*_{xy}, B^*_{xy}) = (A_{xy}, B_{xy})$*; all other entries are set to* $(0, 0)$*. Then run maximin matching on* $\mathbf{P}^{*(n)}$*.*

We wish to determine how the evaluation probability p must vary with n.

Theorem 4 *Let* μ *be the outcome of p-partial maximin matching in a random preference structure* $\mathbf{P}^{(n)}_{\text{ran}}$*. Then*

$$\lim_{n\to\infty} E[\hat{B}_{\mathbf{P}^{(n)}_{\text{ran}}}(\mu)] = 0$$

if and only if $np \to \infty$*.*

Proof That $np \to \infty$ is a necessary condition is almost trivial: Suppose np, the expected number of partners evaluated by any agent, is limited to K. Then any agent would typically prefer n/K unevaluated partners over the best evaluated partner. Hence, even if it would be possible to match every agent with his or her best evaluated partner, there would be a probability $1/K^2$ that a given pair would be blocking.

To prove that $np \to \infty$ is a sufficient condition, we can follow the proof of Theorem 3. The only exception is the last bit of the proof of Lemma 4 where p enters: $\min(A^*_{xy}, B^*_{xy}) \leq (1-\delta)(n-1)$ is at most $1 - p(\delta/2)^2$. The proof then proceeds along the same lines, yielding

$$\begin{aligned} P(D_n) &\leq C(2\pi\epsilon(1-\epsilon)n)^{-1}(1-\epsilon)^{-2(1-\epsilon)n}\epsilon^{-2\epsilon n}(1-p(\delta/2)^2)^{\epsilon^2 n^2} \\ &\leq \exp -[\log(2\pi\epsilon/C) + 2n\log(1-\epsilon) + 2\epsilon n \log\epsilon + \epsilon^2 n^2 p(\delta/2)^2]. \end{aligned}$$

As both n and np tend to infinity, the last term of the exponent will grow faster than linearly, thereby dominating the expression which will hence tend to zero. This proves Lemma 4 for p-partial maximin matching, and the theorem follows.

Thus, the proportion of evaluated partners may decline as n grows.

6 Discussion

Bergstrom and Real (2000) suggested that two-sided stable matching theory can provide insights to the study of animal mate choice. They were dismissed by Simão and Todd (2002, p. 116) who claim that this theory's emphasis on "full knowledge and stability only diverts attention from other issues of greater empirical relevance." Indeed, economic experiments on decentralized matching markets with incomplete information (Ünver, 2005; Niederle and Roth, 2006) have yielded outcomes with considerable instability. However, we believe that these results are not necessarily representative of larger decentralized markets including mate search situations. Our Theorem 4 suggests that if agents put some more effort into mate search than just picking a random partner, and if they increase the search effort somewhat when there are more potential mates, then we can expect outcomes of large markets with a very small proportion of blocking pairs.

References

1. Abraham, D., Biró, P., Manlove, D. F. (2006), "Almost stable" matchings in the roommates problem, in *Proceedings of WAOA 2005*, Springer Lecture Notes in Computer Science vol. 3879, pp. 1–14.
2. Alpern, S., Reyniers, D. (1999) Strategic mating with homotypic preferences, *J. Theo. Biol.* **198**, 71–88.
3. Alpern, S., Reyniers, D. (2005) Strategic mating with common preferences, *J. Theo. Biol.* **237**, 337–354
4. Bergstrom, C. T., Real, L. A. (2000) Towards a theory of mutual mate choice: Lessons from two-sided matching, *Evolutionary Ecology Research* **2**, 493–508.

5. Eriksson, K., Strimling, P. (2005) How unstable are matchings from decentralized mate search?, unpublished manuscript.
Operations Research.

6. Eriksson, K., Sjöstrand, J., Strimling, P. (2006) Optimal expected rank in a two-sided secretary problem, to appear in *Operations Research.*

7. Gale, D., Shapley, L. (1962). College admissions and the stability of marriage, *Amer. Math. Monthly*, **69**: 9–15.

8. Gusfield, D., Irving, R. W. (1989) *The stable marriage problem: Structure and algorithms*, MIT Press.

9. Kagel, J. H., Roth, A. E. (2000) The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment, *Quart. J. Econ.* **115**, 201–235.

10. Kalick, S. M., Hamilton, T. E. (1986) The matching hypothesis reexamined, *Journal of Personality and Social Pscyhology* **51(4)**, 673–682.

11. Niederle, M., Roth, A. E. (2006) Making markets thick: How norms governing exploding offers affect market performance, preprint.

12. Pittel, B. (1989) The average number of stable matchings, *SIAM Journal of Discrete Mathematics* **2**, 530–549.

13. Roth, A. E., Sotomayor, M. A. O. (1990) *Two-sided matching*, Cambridge University Press.

14. Roth, A. E., Xing, X. (1997) Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists, *Journal of Political Economy* **105**, 284–329.

15. Simão, J., Todd, P. M. (2002) Modeling mate choice in monogamous mating systems with courtship, *Adaptive Behavior* **10(2)**, 113–136.

16. Todd, P. M., Miller, G. F. (1999) From pride and prejudice to persuasion: Satisficing in mate search. In G. Gigerenzer, P. M. Todd (Eds.), *Simple heuristics that make us smart*, Oxford University Press, pp. 287–308.

17. Ünver, M. U. (2005) On the survival of some unstable two-sided matching mechanisms, *International Journal of Game Theory*, **33(2)**, 239-254.