University of Connecticut

Department of Economics Working Paper Series

Marriage Matching and the Proposers' Advantage

Onur Celik
University of Connecticut

Vicki Knoblauch
University of Connecticut

Working Paper 2005-30

June 2005

341 Mansfield Road, Unit 1063
Storrs, CT 06269–1063
Phone: (860) 486–3022
Fax: (860) 486–4463
http://www.econ.uconn.edu/

Abstract

Despite an extensive literature exploring two-sided matching problems, there remains much to learn about even the simplest marriage matching model. We adopt as our primary tool a simple measure of how well men do and how well women do under a given matching, and use this tool to demonstrate that a group with randomly generated preferences does very well when matched with a group with identical preferences, and that if both groups. preferences are randomly generated, then the proposers' advantage is quite large. We then use theoretical calculations (that is, we examine randomly generated examples) to illustrate, evaluate and extend our findings.

Journal of Economic Literature Classification: C78, D63, D70

Keywords: Two-Sided Matching, Mechanism, Proposers' Advantage

The authors would like to thank Fusun Yaman for programming assistance.

Introduction.

In this paper we investigate the proposers' advantage in two-sided matching problems. Roth and Peranson (1999) found that for the National Resident Matching Program 1987 and 1993-1996 very few applicants would have benefitted from a change from the program-propose matching algorithm to the applicant-propose matching algorithm, so that there was essentially no proposers' advantage. They ascribe this fact to correlation among applicants' preferences over programs and among proposers' preferences over applicants, and to restrictions on the length of applicants' and programs' preference lists. They also found that even when preferences are randomly generated, if limits are placed on the length of preference lists, the proposers' advantage will be small–see Roth and Peranson (1999, figure 2).

Therefore in order to further our understanding of the proposers' advantage we will focus on the simplest two-sided matching problem, marriage matching with an equal number n of men and women with complete and transitive strict preferences truthfully reported such that no man finds any woman unacceptable and no woman finds any man unacceptable. For much of this study we will assume men's and women's preferences are randomly generated.

We will first need to adopt a measure of the proposers' advantage. Given men's and women's preferences, for each matching μ of men with women we assign a score $\mu m_i S$ to each man and a score $\mu w_j S$ to each women. We define $\mu m_i S$ to be m_i's ranking of his match under μ and $\mu w_j S$ to be w_j's ranking of her match under μ. We then define the men's score μMS to be the sum of the men's scores, and the women's score μWS to be the sum of the women's scores. We have chosen these definitions from among many possible alternatives such as the sum of the squares of the women's scores if one wants a measure that emphasizes the least successful women, or the sum of the logs of the women's scores if one wants a measure that emphasizes the most successful women. Our choice of the simplest and easiest to work with definitions was dictated by our goal: we want the men's score and the women's score to be effective tools in the investigation of the proposers' advantage and other properties of the men-propose and women-propose matchings μ_M and

μ_W, which will be defined in Section 2. Once obtained, our results can be translated into statements about other versions of men's score, women's score and proposers' advantage.

Finally, we formulate measures of the proposers' advantage: $PAM = \mu_W MS - \mu_M MS$ and $PAW = \mu_M WS - \mu_W WS$ which for given preferences measure how much better a group does when it proposes then when it is proposed to. When men's and women's preferences are randomly generated, PAM and PAW have identical expected values.

Theoretical results established in Section 3 include the following. Men's preferences identical and women's preferences identical is a worst case scenario in that it maximizes the sum of the men's and women's scores (Corollary 1 and Proposition 2). A group with randomly generated preference does very well when matched with a group with identical preferences (Proposition 4). If men's preferences and women's preferences are randomly generated, then the expected value of the proposers' advantage is quite large, at least $\frac{n^2+n}{4\ln n+4} - n(1+\ln n) \sim n^2/4\ln n$ (Corollary 3).

In Section 4 we carry out what Roth and Peranson (1999) have called "theoretical calculations." For several values of n we give men identical ROLs and generate random ROLs for women. Then we calculate the average and standard deviation over 300 trials of the men's score, the women's score and the proposers' advantage. We make the same calculations when all ROLs are randomly generated. The results are presented and compared with our theoretical results in five tables. The theoretical calculations indicate that our theoretical bounds on men's and women's scores are quite tight with the exception of our lower bound on a woman's expected ranking of her men-propose match (all ROLs random). It follows that our lower bound on the proposers' advantage (all ROLs random) is also not tight.

Both Roth and Peranson (1999) and Feldin (1999) discuss n_S the number of stable matchings, and n_M, the number of men matched differently under μ_M and μ_W, when preferences are randomly generated. Irving and Leather (1986) also discuss n_S. There are some similarities with our paper since n_S and n_M can both be considered crude measures of the proposers' advantage. However, Irving and Leather's result, that the maximum value of n_S is at least 2^{n-1}, doesn't translate into a significant statement about the proposers' advantage; and Roth and Peranson, and Feldin are interested in using theoretical calculations and theory to show that if the length of men's preference lists is held constant (a nod to bounded rationality), and each woman holds preferences only over those men on whose preference lists she appears, then as the number of men and women is allowed to grow, the expected values of n_S and n_M remain bounded. We investigate a much different scenario, one in which each man ranks all women and vice versa. Our investigation yields new information about the proposers' advantage, augmenting and contrasting with the results of Roth and Peranson and Feldin, who demonstrated that in the short preference-list scenario there is little or no proposers' advantage.

Most other papers in the two-sided matching literature focus on problems in more complex real-world markets, for example Roth, Sönmez and Ünver (2004) and Chen and Sömnez (2002); or on strategic considerations, that is strategic reporting of ROLs, for example Roth and Vande Vate (1991) and Demange, Gale and Sotomayor (1987).

2. Preliminaries.

For the reasons given in the introduction, we will be studying the simplest marriage matching problem. There are n men $\{m_1, m_2, \ldots, m_n\}$ and n women $\{w_1, w_2, \ldots, w_n\}$. Each woman has a ranking order list (ROL) which gives her complete, transitive strict preferences over the n men. The symmetric statement holds for each man. A *matching* is a 1-1 function μ: $\{m_1, m_2, \ldots, m_n\} \cup \{w_1, w_2, \ldots, w_n\} \rightarrow \{m_1, m_2, \ldots, m_n\} \cup \{w_1, w_2, \ldots, w_n\}$ such that $\mu(m_i) \in \{w_1, w_2, \ldots, w_n\}$ for all i and $\mu(w_j) \in \{m_1, m_2, \ldots, m_n\}$ for all j. In other words, a matching is a collection of n ordered pairs $< m_i, w_j >$ such that each man and each woman occurs in exactly one ordered pair. A matching μ is *unstable* if there exits a man m_i and a woman w_j such that m_i prefers w_j to $\mu(m_i)$ and w_j prefers m_i to $\mu(w_j)$. A matching is *stable* is it is not unstable.

The *men-propose matching* μ_M is defined as follows. In round 1 each man proposes to his top-ranked woman. Each woman proposed to tentatively accepts her top-ranked man among those who proposed to her. In round 2 each man who is not engaged proposes to his top ranked woman among those to whom he has not yet proposed. Each woman proposed to tentatively accepts her top ranked man among those who proposed to her in round 2 and the man to whom she is engaged (if she is engaged). The procedure continues until all men (and therefore all women) are engaged, at which time the tentative matches become final. The resulting matching is the men-propose matching μ_M. The *women-propose matching* μ_W is defined symmetrically.

Both μ_M and μ_W are stable matchings (Gale and Shapley, 1962). The matching μ_W is women optimal (Gale and Shapley, 1962); that is, for each j $\mu_W(w_j)$ is w_j's top-ranked man among all men with whom she is matched by some stable matching. Also, μ_W is men worst in that for all i $\mu_W(m_i)$ is m_i's bottom-ranked woman among all women with whom he is matched by some stable matching. The symmetric statements hold for μ_M.

For any matching μ, m_i's score $\mu m_i S$ is m_i's ranking of $\mu(m_i)$. The men's score μMS is defined by $\mu MS = \sum_{i=1}^{n} \mu m_i S$. The women's score is defined similarly.

The measures we will use for the proposer's advantage are $PAM = \mu_W MS - \mu_M MS$, and $PAW = \mu_M WS - \mu_W WS$, both of which are non-negative since μ_M is men optimal and μ_W is women optimal. These measures tell us for fixed preferences how much better a group does when it proposes than when it is proposed to.

3. Bounds on the Proposers' Advantage.

3.1 Men's ROLs Identical.

Proposition 1 establishes some bounds on men's and women's scores when men's ROLs are perfectly correlated.

Proposition 1. *If the men's ROLs are identical, then*

1) $\mu_M MS = \mu_W MS = (n^2 + n)/2$.
2) $\mu_M = \mu_W$
3) $n \leq \mu_M WS = \mu_W WS \leq (n^2 + n)/2$ *and the bounds are attainable.*
4) $PAM = PAW = 0$,

Proof.

1) In each of μ_M and μ_W some man is matched with the top-ranked woman for a score of +1; some man is matched with the second ranked woman for a score of +2; etc.

$$\mu_M MS = \mu_W MS = 1 + 2 + \ldots + n = (n^2 + n)/2$$

2) Since each man does at least as well in μ_M as in μ_W, by 1) each man does exactly as well. In other words, $\mu_M = \mu_W$.

3) Clearly $n \leq \mu_M WS$. By 2) $\mu_M WS = \mu_W WS$.

In round 1 of μ_M, the top-ranked woman is matched with her top-ranked man; in round 2 the second-ranked woman is matched with her top-ranked or second-ranked man (depending on whether her top-ranked man was taken by the top-ranked woman); in round 3 the third-ranked woman is matched with her top-ranked or second-ranked or third-ranked man; etc. Therefore

$\mu_M WS \leq 1 + 2 \ldots + n = (n^2 + n)/2$

If no two women rank the same man first $\mu_W WS = n$.

If the women's ROLs are identical, then 1) restated for women yields $\mu_M WS = (n^2 + n)/2$

4) $PAM = \mu_W MS - \mu_M MS = 0$ by 1)

$PAW = \mu_M WS - \mu_W WS = 0$ by 3). ∎

Corollary 1. *If the men's ROLs are identical, and the women's ROLs are identical and μ is any stable matching, then*

$$\mu MS + \mu WS = n^2 + n$$

Proof. By Gale and Shapley (1962) μ_M and μ_W give each man, respectively, his best and worst matches over all stable matchings. Therefore, by 1) of Proposition 1 $\mu MS = (n^2 + n)/2$. Since the women's ROLs are identical, a symmetric argument gives $\mu WS = (n^2 + n)/2$. ∎

Corollary 1 and the following proposition show that men's ROLs identical and women's ROLs identical is a worst-case scenario in that it maximizes the sum of the men's score and the women's score.

Proposition 2. *For any ROLs and any stable matching* μ

$$\mu MS + \mu WS \leq n^2 + n$$

Proof. Since μ is stable, for each ordered pair (m_i, w_j) not matched by μ, at most one of "m_i prefers w_j to $\mu(m_i)$" and "w_j prefers m_i to $\mu(w_j)$" is true. Since there are $n^2 - n$ unmatched pairs

$$\mu MS + \mu WS \leq n^2 - n + 2n = n^2 + n \quad \blacksquare$$

We found in Proposition 1 that when men's ROLs are identical, the men's score is $(n^2 + n)/2$ and the women's score can be as low as n. Proposition 3 shows that this is not a worst case scenario for men vs. women in that for some ROLs when men propose the women's score can be n and the men's score much greater than $(n^2 + n)/2$.

Proposition 3. $\max(\mu_M MS - \mu_M WS) = n^2 - 2n + 1$, where the maximum is taken over all ROLs.

Proof. First we show $\mu_M MS \leq n^2 - (n - 1)$ for any ROLs. In the last round of proposals, at least one woman is proposed to for the first time. Let w_a be one such woman. Each of the $n - 1$ men not matched with w_a never proposed to w_a Therefore, w_a is ranked lower by each of them than the woman to whom he is matched. Therefore, $\mu_M MS \leq n^2 - (n - 1)$. Since $\mu_M WS \geq n$,

$$\mu_M MS - \mu_M WS \leq (n^2 - n + 1) - n = n^2 - 2n + 1$$

Next we show that this bound is achievable. Suppose the ROLs are

m_1: $w_1, w_2, \ldots, w_n$

m_i: $w_i, w_{i+1}, \ldots, w_{n-1}, w_1, w_2, \ldots, w_{i-1}, w_n$ for $2 \leq i \leq n - 1$

m_n: $w_{n-1}, w_1, w_2, \ldots, w_{n-2}, w_n$

w_i: $m_{i+1}, m_{i+2}, \ldots, m_n, m_1, m_2, \ldots, m_i$ for $1 \leq i \leq n-2$

w_{n-1}: $m_1, m_2, \ldots, m_n$

w_n: $m_n, m_1, m_2, \ldots, m_{n-1}$

The reader can verify that μ_M applied to these ROLs takes $(n-1)^2+1$ rounds for a total of $n^2-(n-1)$ proposals and that $\mu_M MS = n^2-(n-1)$ and $\mu_M WS = n$.

3.2 Men's ROLs identical and women's ROLs randomly generated.

Now we will show that randomly chosen ROLs do very well against perfectly correlated ROLs. Since women's ROL's are randomly chosen we will be dealing with the expected values $E(.)$ of men's and women's scores.

Proposition 4. *If the men's ROLs are identical and the women's ROLs are randomly chosen, then*

$$E(\mu_M MS) = E(\mu_W MS) = (n^2+n)/2$$

and

$$(n+1)(\ln(n+1)-\ln 2) \leq E(\mu_M WS) = E(\mu_W WS) \leq (n+1)(\ln(n+1)-\ln 2)+(n+1)/2$$

Proof. The following lemma will be used in the proof.

Lemma 1. *If $X_n = \{1,2,\ldots,n\}$, $S_{k,n} \subseteq X_n$ is a randomly chosen subset of cardinality k and $1 \leq k \leq n$, then*

$$E(\min S_{k,n}) = \frac{n+1}{k+1}$$

Proof of Lemma 1. The proof proceeds by induction on n. If $n=1$, then $k=1$ and $E(\min S_{1,1}) = 1 = \frac{n+1}{k+1}$.

Next we show that if the conclusion holds for some positive integer N, then it holds for $N+1$.

Induction Hypothesis. If $1 \leq k \leq N$, then $E(\min S_{k,N}) = \frac{N+1}{k+1}$

If $1 \leq k \leq N+1$, then

$E(\min S_{k,N+1}) = (\frac{k}{N+1})1 + \frac{(N+1)-k}{N+1}\ (1+\frac{N+1}{k+1})$

since $\frac{k}{N+1}$ is the probability that $1 \in S_{k,N+1}$, $\frac{(N+1)-k}{N+1}$ is the probability that $1 \notin S_{k,N+1}$ and by the induction hypothesis $\frac{N+1}{k+1}$ is the expected minimum of a randomly chosen subset of $\{1, 2, \ldots N\}$ of cardinality k so that $1 + \frac{N+1}{k+1}$ is the expected minimum of a randomly chosen subset of $\{2, 3, \ldots, N+1\}$ of cardinality k.

Simplifying,

$E(\min S_{k,N+1}) = \frac{N+2}{k+1}$. ■

Returning to the proof of Proposition 4, the equalities in the statement of the Proposition follow from Proposition 1.

Without loss of generality, assume w_j is the j^{th} ranked woman for $1 \leq j \leq n$. Then

$$\begin{aligned} E(\mu_M WS) &= \sum_{j=1}^{n} E(\mu_M w_j S) \\ &= \sum_{j=1}^{n} E(\min(S_{n-j+1}, n)) \end{aligned}$$

since $w_j's$ score is the minimum of the ranks of the $n - j + 1$ men who propose to her in round j. Using Lemma 1

$$\sum_{j=1}^{n} E(\min(S_{n-j+1}, n)) = \sum_{k=1}^{n} E(\min S_{k,n}) = \sum_{k=1}^{n} \frac{n+1}{k+1}$$

One can see on a graph of $f(x) = \frac{n+1}{x+1}$ that

$$\int_1^n \frac{n+1}{x+1} dx \leq \sum_{k=1}^{n} \frac{n+1}{k+1} \leq \int_1^n \frac{n+1}{x+1} dx + (n+1)/2$$

Therefore

$$(n+1)(\ln(n+1) - \ln 2) \leq E(\mu_M WS) \leq (n+1)(\ln(n+1) - \ln 2) + (n+1)/2 \quad ■$$

Proposition 4 tells us that for large n random women's ROLs put the expected value of $\mu_M WS$ much nearer to the lower bound n in Proposition 1 conclusion 3) than to the upper bound $(n^2 + n)/2$. For example for $n = 10$ $\quad 10 \le 18 \le E(\mu_M WS) \le 25 \le 55$

for $n = 100$ $\quad 100 \le 396 \le E(\mu_M WS) \le 447 \le 5,050$

for $n = 1000$ $\quad 1,000 \le 6,221 \le E(\mu_M WS) \le 6,723 \le 500,500$

3.3 Men's and Women's ROLs Randomly Generated.

We begin with a lemma for numbered balls drawn from an urn.

Lemma 2. *For $1 \le k \le n$, the number $D_{n,k}$ of random draws with replacement from $\{1, 2, \ldots, n\}$ until every element of $\{1, 2, \ldots, k\}$ has been drawn at least once satisfies*

$$E(D_{n,k}) = n + n/2 + n/3 + \ldots + n/k \le n(1 + \ln k) \tag{1}$$

and

$$Prob\ (D_{n,k} \le n(1 + 2\ln k)) \ge 1/2 \tag{2}$$

Proof. Define $E_{n,0} = 0$. For $1 \le k \le n$ write $E_{n,k}$ for $E(D_{n,k})$. Since k/n is the probability that the first draw comes from $\{1, 2, \ldots, k\}$

$$E_{n,k} = \tfrac{k}{n}(1 + E_{n,k-1}) + \tfrac{n-k}{n}(1 + E_{n,k})$$

Simplifying,

$$E_{n,k} = E_{n,k-1} + n/k$$

Solving recursively,

$$\begin{aligned} E_{n,k} &= n + n/2 + n/3 + \ldots + n/k \\ &\le n(1 + \textstyle\int_1^k \frac{dx}{x}) = n(1 + \ln k) \end{aligned}$$

This establishes (1).

If $Prob\ (D_{n,k} \le n(1 + 2\ln k)) < 1/2$, then $E(D_{n,k}) > n(1 + \ln k)$, contradicting (1). This establishes (2). ∎

Now we apply Lemma 2 to establish bounds on the expected number of proposals under the men-propose matching algorithm.

Proposition 5. *For n men and n women, all ROLs randomly generated,*

$$E(\#\mu_M \textit{ proposals}) \ \leq n(1+\ln n) \tag{3}$$

and

$$\textit{Prob}\ (\#\mu_M \textit{ proposals} \ \leq n(1+2\ln n)) \geq 1/2 \tag{4}$$

Proof. First, $E(\#\mu_M \textit{ proposals}) \leq E(D_{n,n})$, since each proposal is a random draw with only partial replacements, (in each round, the proposals can be thought of as occurring sequentially and since ROLs are random, each man who proposes draws randomly not from all of $\{w_1, w_2, \ldots, w_n\}$ but only from the set of women to whom he has not yet proposed). Similarly, $\textit{Prob}(\#\mu_M \textit{ proposals} \leq n(1+2\ln n)) \geq \textit{Prob}\ (D_{n,n} \leq n(1+2\ln n))$. Therefore (3) and (4) follow from Lemma 2. ▪

Proposition 6. *For n men and n women, all ROLs randomly generated,*

$$E(\mu_M MS) \leq n(1+\ln n) \tag{5}$$

and

$$E(\mu_M WS) \geq \frac{n^2+n}{4\ln n+4} \sim n^2/4\ln n \tag{6}$$

Proof. Inequality (5) follows from (3) of Proposition 5 and the fact that $\mu_M MS$ is equal to the number of proposals made.

Now suppose k is a positive integer, $x_1, x_2, \ldots, x_n$ are non-negative integers and $\sum_{j=1}^n x_j = k$. What can we say about $E(\mu_M WS)$ given that for $1 \leq j \leq n$, w_j is proposed to exactly x_j times?

$E(\mu_M WS\colon\ x_1, x_2, \ldots, x_n, k) = \sum_{j=1}^n E(\mu_M w_j S)$
$= \sum_{j=1}^n \frac{n+1}{x_j+1}$ by Lemma 1.

Lemma 3. *Given non-negative reals $x_1, x_2, \ldots x_n$ with $\sum_{j=1}^n x_j = k > 0$, $\sum_{j=1}^n \frac{n+1}{x_j+1} \geq n\frac{n+1}{k/n+1}$*

Proof. The function $f(x_1, \ldots, x_n) = \sum_{j=1}^n \frac{n+1}{x_j+1}$ has a minimum on $X = \{(x_1, \ldots, x_n) \in \Re^{n+}\colon\ \sum_{j=1}^n x_j = k\}$.

For $(x_1, \ldots, x_n) \in X$ with $x_b \neq x_c$,

$$\sum_{j=1}^{n} \frac{n+1}{x_j+1} = \sum_{\substack{j=1 \\ j\neq b,c}}^{n} \frac{n+1}{x_j+1} + \frac{n+1}{x_b+1} + \frac{n+1}{x_c+1}$$

$$\geq \sum_{\substack{j=1 \\ j\neq b,c}}^{n} \frac{n+1}{x_j+1} + \frac{n+1}{(x_b+x_c)/2+1} + \frac{n+1}{(x_b+x_c)/2+1}$$

Therefore the minimum of f can't occur when two x^j's are unequal. The minimum occurs when $x_j = k/n$ for all j. ∎

$E(\mu_M WS) =$

$$\sum_{k=n}^{n^2-n+1} \sum_{\substack{x_1,x_2,\ldots,x_n \text{with} \\ x_1+x_2+\ldots x_j=k}} \text{Prob } (\#\mu_M \text{ proposals to } w_j \text{ is } x_j \text{ for } 1 \leq j \leq n)\times$$

$$E(\mu_M WS\text{: } \#\mu_M \text{ proposals to } w_j \text{ is } x_j \text{ for } 1 \leq j \leq n)$$

and for k fixed

$$E(\mu_M WS\text{: } \#\mu_M \text{ proposals to } w_j = x_j \text{ for } 1 \leq j \leq n) =$$

$$\sum_{j=1}^{n} E(\mu_M w_j S\text{: } \#\mu_M \text{ proposals to } w_j \text{ is } x_j) =$$

$$\sum_{j=1}^{n} \frac{n+1}{x_j+1} \text{ by Lemma 1}$$

$$\geq n \frac{n+1}{k/n+1} \text{ by Lemma 3}$$

Therefore, with [.] representing the greatest integer function,

$$E(\mu_M WS) \geq \sum_{k=n}^{[n(1+2\ln n)]} \text{Prob } (\#\mu_M \text{ proposals} = k)\ n \frac{n+1}{k/n+1}$$

$$\geq \text{Prob } (\#\mu_M \text{ proposals } \leq n(1+2\ln n))\ n\ \frac{n+1}{(1+2\ln n)+1}$$

Applying (4) to the first factor,

$$E(\mu_M WS) \geq \frac{n^2+n}{4\ln n+4} \sim n^2/4\ln n. \quad ∎$$

Proposition 6 can be restated in terms of individuals:

Corollary 2. *If there are n men and n women and ROLs are randomly generated, then each man's expected ranking of his μ_M match is less than or equal to $1 + \ln n$ and each woman's expected ranking of her μ_M match is greater than $\frac{n+1}{4 \ln n+4} \sim n/4 \ln n$.*

By the symmetry of the assumption that men's and women's ROLs are randomly generated, $E(\mu_M MS) = E(\mu_W WS)$ and $E(\mu_M WS) = E(\mu_W MS)$. Therefore $E(PAM) = E(PAW)$. Proposition 6 provides the following lower bound on the proposers' advantage.

Corollary 3. *For n men and n women with ROLs generated randomly,* $E(PAW) = E(PAM) = E(\mu_W MS) - E(\mu_M MS) \geq \frac{n^2+n}{4 \ln n+4} - n(1 + \ln n) \sim n^2/4 \ln n$.

4. Theoretical Calculations.

In our first set of calculations, for each value of $n = 3, 5, 10, 25, 50, 100, 500, 1000$ we let $w_1, w_2, \ldots, w_n$ be the ROL of every man, and we generated the women's ROLs using a pseudo random number generator. Then we implemented the men-propose matching μ_M. We calculated the average and standard deviation over 300 trials of the men's score and the women's score. The results are given in Table 1 which can be found in the Appendix. It can be seen that $\mu_M MS = \mu_W WS = (n^2 + n)/2$ per Proposition 1. It can also be seen by comparison of the two columns in Table 1 that under μ_M the proposees do much better on average than the proposers, and that the proposees' score has a fairly high standard deviation.

Table 2 compares the 300 trial averages with the theoretical lower and upper bounds from Proposition 4, $\underline{E}(\mu_M WS) = (n+1)(\ln(n+1) - \ln 2)$ and $\overline{E}(\mu_M WS) = (n+1)(\ln(n+1) - \ln 2) + (n+1)/2$. The theoretical bounds are seen to be quite tight.

In our second set of calculations, for each value of $n = 3, 5, 10, 25, 50, 100, 500, 1000$ we generated all ROLs using a pseudo random number generator. Then we implemented μ_M and μ_W. We calculated the average and standard deviation over 300 trials of the men's score and the women's score. The results are given in Table 3.

The first two columns of Table 4 compare the average ranking by a man of his μ_M

match (equal to the corresponding entry from Table 3 divided by n) with $\overline{E}(\mu_M m_i S) = 1 + \ln n$, the theoretical upper bound from Corollary 2 on each man's expected ranking of his μ_M match. It can be seen that this bound is quite tight.

We display individual scores in Table 4 rather than group scores because individual scores give a better feel for how well men do under μ_M when ROLs are random; for example, each of 500 men expects his 7th-ranked woman. (Individual scores were not used in Table 2 because they would be somewhat misleading there since with men's lists identical and women's lists random, women have different expected rankings for their μ_M matches; for example, the first ranked woman is matched with her first-ranked man, while the nth-ranked woman expects her $(n+1)/2$-ranked man.)

The last two columns of Table 4 compare the average ranking by a woman of her μ_M match with $\underline{E}(\mu_M w_j S) = (n+1)/(4\ln n + 4)$, the theoretical lower bound from Corollary 2 on each woman's expected ranking of her μ_M match. This bound is quite loose, but at $n = 500$ it begins to yield some information, since it is there that it first exceeds $\overline{E}(\mu_M m_i S)$.

Table 5 compares the theoretical lower bound on the proposer's advantage from Corollary 3 with the average proposer's advantage, AvePAM, which can be calculated by subtracting column 1 of Table 3 from column 4 of Table 3 and with AvePAW, which can be calculated by subtracting column 3 of Table 3 from column 2 of Table 3. The theoretical lower bound only begins to yield information at $n = 500$, since it is there that it becomes positive, but it quickly becomes quite large since it is on the order of $n^2/\ln n$.

5. Concluding Remarks.

We have seen that when one group's ROLs are identical, there is no proposers' advantage; when men's ROLs are identical and women's ROLs random, women do much better than men; and when all ROLs are random, there is a large proposers' advantage.

There remain many interesting questions related to simple marriage matching; for example, do successful men marry successful women; that is, what is the relationship between a man's ranking *of* his match and his ranking *by* his match? There are ROLs and stable matchings for which a man's score is positively correlated with his match's score, and there are ROLs and stable matchings for which the correlation is negative. Is there more that can be said in answer to this question?

References

1. Chen, Yan and Tayfun Sönmez. "Improving Efficiency of On-Campus Housing: An Experimental Study," *American Economic Review* **92** (2002), 1669-1686.
2. Demange, Gabrielle, David Gale and Marilda Sotomayor. "A Further Note on the Stable Matching Problem," *Discrete Applied Mathematics* **16** (1987), 217-22.
3. Feldin, Aljosa. "Core Convergence in Two-Sided Matching Markets: Some Theoretical Considerations," Mimeo. University of Pittsburgh (1999).
4. Gale, David and Lloyd S. Shapley. "College Admissions and the Stability of Marriage," *American Mathematical Monthly* **69** (1962), 9-15.
5. Irving, Robert, W. and Paul Leather. "The Complexity of Counting Stable Marriages," *SIAM Journal of Computing* **15** (1986), 655-667.
6. Roth, Alvin E. and John H. Vande Vate. "Incentives in Two-Sided Matching with Random Stable Mechanisms,"*Economic Theory* **1** (1991), 31-44.
7. Roth, Alvin E. and Elliott Peranson. "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *The American Economic Review* **89** (1999), 748-780.
8. Roth, Alvin E., Tayfun Sönmez and M. Utku Ünver. "Kidney Exchange," *Quarterly Journal of Economics* **19** (2004), 457-488.

Appendix

Table 1. Men's ROLs identical, Women's ROLs random
(300 observations for each n $\leq$ 500, 40 observations for n=1,000)

n		μ_M MS (= μ_W MS)	μ_M WS (= μ_W WS)
3	Ave.	6	4.39
	St. D.	0	.913
5	Ave.	15	8.49
	St. D.	0	2.05
10	Ave.	55	22.2
	St D.	0	4.52
25	Ave.	325	73.9
	St D.	0	13.3
50	Ave.	1,275	181
	St. D.	0	28.1
100	Ave.	5,050	423
	St. D.	0	61.5
500	Ave.	125,250	2,910
	St. D.	0	292
1,000	Ave.	500,500	6,260
	St. D.	0	536

Table 2. Men's ROLs identical, Women's ROLs random
Comparison of theoretical computations (300 observations for each n $\leq$ 500, 40 observations for n=1,000) with theoretical bounds from Proposition 4.

n	$\underline{E}(\mu_M \text{WS})$	Ave μ_MWS	$\overline{E}(\mu_M \text{WS})$
3	2.77	4.39	4.77
5	6.59	8.49	9.59
10	18.8	22.2	24.3
25	66.7	73.9	79.7
50	165	181	191
100	396	423	447
500	2,770	2,910	3,020
1,000	6,220	6,260	6,720

Table 3. All ROLs random.
(300 observations for each n $\leq$ 500, 40 observations for n=1,000)

n		μ_MMS	μ_MWS	μ_WWS	μ_WMS
3	Ave.	4.30	5.27	4.55	5.02
	St. D.	.966	1.28	1.04	1.35
5	Ave.	8.71	12.2	8.67	12.3
	St. D.	2.33	3.37	2.38	3.06
10	Ave.	23.9	36.8	24.0	36.6
	St D.	6.04	9.06	6.56	9.47
25	Ave.	86.1	174	85.4	175
	St D.	21.1	39.8	21.7	39.7
50	Ave.	207	592	210	579
	St. D.	48.4	121	48.4	129
100	Ave.	509	1,970	496	2,020
	St. D.	112	378	136	396
500	Ave.	3,460	37,000	3,330	38,300
	St. D.	685	6,540	614	6,340
1,000	Ave.	6,880	149,000	7,550	134,000
	St. D.	1,360	24,000	731	14,100

Table 4. All ROLs random.
Comparison of theoretical calculations (300 observations for each n $\le$ 500, 40 observations for n=1,000) with theoretical bounds from Corollary 2.

n	Ave(μ_M m_i S) Ave(μ_W w_j S)	$\overline{E}$(μ_M m_i S)		$\underline{E}$(μ_M w_j S)	Ave(μ_M w_j S) Ave(μ_W m_i S)
3	1.43 1.52	2.10		.477	1.76 1.67
5	1.74 1.73	2.61		.575	2.44 2.46
10	2.39 2.40	3.30		.833	3.68 3.66
25	3.44 3.42	4.22		1.54	6.96 7.00
50	4.14 4.20	4.91		2.60	11.8 11.6
100	5.09 4.96	5.61		4.50	19.7 20.2
500	6.92 6.66	7.21		17.4	74.0 76.6
1,000	6.88 7.55	7.91		31.6	149 134

Table 5. All ROLs random.
Comparison of average proposers' advantage (300 observations for each $n \leq 500$, 40 observations for n=1,000) with $\underline{E}$ *(PAM)*, the theoretical lower bound on proposers' advantage from Corollary 3.

n	$\underline{E}$ (*PAM*)	Ave *PAM*	Ave *PAW*
3	0	.720	.720
5	0	3.59	3.53
10	0	12.7	12.8
25	0	88.9	88.6
50	0	372	382
100	0	1,510	1,470
500	5,070	34,800	33,700
1,000	23,700	127,000	141,000