Marriage Matching and Intercorrelation of Preferences

James W. Boudreau
University of Connecticut

Vicki Knoblauch
University of Connecticut

Abstract

Men's and women's preferences are intercorrelated to the extent that men rank highly those women who rank them highly. Intercorrelation plays an important but overlooked role in determining outcomes of matching mechanisms. We study via simulation the effect of intercorrelated preferences on men's and women's aggregate satisfaction with the outcome of the Gale-Shapley matching mechanism. We conclude with an application of our results to the student admission matching problem.

1 Introduction.

In two-sided matching problems, the intercorrelation of preferences plays an important but largely overlooked role in determining the outcomes of various matching mechanisms. Other factors that also play important roles include correlated preferences, truncated preference lists and strategic reporting of preferences. These better known factors are defined in terms of a single side's preferences, or, since we will concentrate on marriage matching, in terms of a single gender's preferences. In contrast, intercorrelation of preferences concerns the *relationship* between men's and women's preferences: roughly speaking, men's and women's preferences are intercorrelated to the extent that men rank highly those women who rank them highly.

Our main goal in this paper is a quantitative description of the effect of intercorrelation of preferences on men's and women's aggregate satisfaction with the outcome of the Gale-Shapley matching algorithm. We set this as our goal in the belief that, or rather with the knowledge that, while there is a great deal to be learned about those complex two-sided matching problems that arise in real-life applications, there is still much to be learned about two-sided matching by studying the workings of the Gale-Shapley algorithm applied to the simplest marriage matching problem.

We choose to use simulation in our study because theoretical progress is difficult in this area due to the complex interaction of men's and women's preferences during a run of the men-propose Gale-Shapley algorithm. Theoretical progress has been slow and known results involve unrealistic and restrictive assumptions about preferences. For example, Wilson (1972) found a useful upper bound for the expected sum of men's rankings of their mates as assigned by the men-propose Gale-Shapley algorithm when women's preferences are arbitrary (unrealistic) and men's are random (unrealistic and restrictive). Knoblauch (2008) found a lower bound which, together with Wilson's upper bound, implies the expected sum of men's rankings of their assigned mates divided by the natural log of the number of men approaches 1 for large n. Using simulation, we will be able to go beyond the assumptions of women's preferences arbitrary and men's random and investigate more realistic cases in which men's preferences are to various degrees correlated, as are women's, and men's and women's preferences are to various degrees intercorrelated.

Before outlining our plan of attack, we present three examples to show how positive and negative intercorrelation of preferences arise.

Example 1. Positively Intercorrelated Preferences. Chen and Sönmez (2006) study student assignment in Boston, where law mandates that schools give preference to students who live nearby and/or have siblings already enrolled. Presumably, students prefer nearby schools and schools their siblings attend. The result is positive intercorrelation of preferences.

While student assignment in Boston is a many-to-one matching problem, it can be converted to a one-to-one matching problem by representing each school as many individuals with identical preferences (Teo, Sethuraman and Tan, 2001). This replication procedure maintains positive intercorrelation of preferences while introducing some correlation, since replicants of the same school have identical preferences.

Example 2. Positively Intercorrelated Preferences. Consider a world-wide pen-pal club that matches children with opposite gender pen-pals. In this example, individuals will prefer not the nearby matches as in Example 1, but the more exotic distant matches. The result will again be positively intercorrelated preferences.

Example 3. Negatively Intercorrelated Preferences. After public universities have filled their mandated in-state student quotas, there remains the problem of matching unfilled slots with unmatched applicants. Due to higher out-of-state tuition, universities will tend to prefer out-of-state applicants, while applicants will tend to prefer in-state universities. The result will be negatively intercorrelated preferences.

We approach the question of the effect of intercorrelated preferences on the outcome of the men-propose G-S algorithm as follows. First we develop a method for generating men's and women's preferences with arbitrary levels of correlation and intercorrelation. Then we run the men-propose G-S algorithm on many marriage matching problems that have been generated by our method. Next we run regressions on aggregate satisfaction vs. intercorrelation of preferences. Of course we find as expected that aggregate satisfaction increases for both genders with increasing intercorrelation of preferences, but we analyze our results more closely in order to disentangle the effects of intercorrelation from those of correlation. We also develop a general measure of intercorrelation; the measure used in the aforementioned analysis is only defined for preferences generated by our method.

We should remark here that we are not trying to define the best or most natural measure of intercorrelation. Rather we are looking for a measure of intercorrelation that works. In other words, we want a measure of intercorrelation that allows us to demonstrate that, with just three numbers, correlation of men's lists and of women's lists and intercorrelation, we can to some extent predict the outcome of the complex interaction of men's and women's preferences that occurs during a run of the Gale-Shapley algorithm.

Finally, we discuss the effect of intercorrelation on the number of women who would profit by a shift from the men-propose to women-propose G-S algorithm, which is a proxy for the number of stable matchings, and which in turn is a proxy for the susceptibility of a matching mechanism to strategic reporting of preferences. For the sake of simplicity and clarity of presentation, we restrict our discussion to marriage matching with 100 men and 100 women, and almost exclusively to the men-propose G-S algorithm. We also concentrate on aggregate gender satisfaction because the anticipated aggregate satisfaction with the outcome of a matching mechanism strongly influences whether a gender will agree to the use of that mechanism.

Even before Gale and Shapley (1962) presented the world with their simple yet powerful deferred acceptance algorithm, a similar centralized matching mechanism was already being used in the market for medical interns. Though controversial at times (and undergoing some adaptive transformations to keep up with market conditions), the National Residency Matching Program (NRMP) has long stood as evidence that G-S matching mechanisms can in fact deliver Pareto efficient outcomes in real world two-sided markets. The evolution of the NRMP, the basic theoretical properties of the G-S algorithm, and many examples and extensions of two-sided matching mechanisms are chronicled by Roth and Sotomayor (1990).

In a thorough account of the NRMP's latest incarnation, Roth and Peranson (1999) acknowledge the crucial role of preferences in determining market outcomes, and emphasize the need for a more in depth design literature to aid future market engineering efforts. Motivated by this call to action, we set out to study intercorrelation of preferences, which, immediately upon its discovery, we had suspected of having a strong influence on matching outcomes. Real world examples of beneficial economic design in two-sided matching environments are plentiful, most notably in public education markets where intercorrelated preferences are undoubtedly present due to location considerations (Teo et al., 2001; Chen and Sonmez, 2001;Abdulkadíroğlu et al.,

2005a and 2005b). Potential for improved matching also exists in employment markets such as the market for law clerks (Haruvy et al. 2006), and in less traditional markets such as the sorority rush market (Mongell and Roth, 1991).

All of the aforementioned examples and the corresponding literature underline the effectiveness of theoretical and computational work in studying the rules that govern market interaction. While G-S matching mechanisms undeniably provide Pareto efficient outcomes in two-sided markets, results regarding satisfaction and strategic possibilities are crucial to elicit full participation and truthful representation. Like Knoblauch and Celik's work (2007) with *correlated* preferences, our work with intercorrelated preferences quantifies the relationship between a key aspect of preference structure and the satisfaction of each side of the market with the G-S outcome. This requires as a key step disentangling the effects of correlation and intercorrelation. As an additional contribution, our work also quantifies the relationship between preference structure and the potential for strategic manipulation. Previous work has shown that a single-valued core is a necessity to truly strategy-proof matching mechanisms so that no individual can benefit from misrepresentation (Sönmez, 1999; Ehlers and Masso, 2005). We quantify the link between a population's preferences and the size of the stable set.

The rest of the paper is organized as follows. Section 2 contains preliminaries. Section 3 introduces a method for generating intercorrelated preferences. Sections 4-7 discuss the effect of intercorrelated preferences on the outcomes of matching mechanisms. Section 8 concludes.

2 Preliminaries.

The simplest marriage matching problem starts with a set $M = \{m_1, m_2, \ldots, m_n\}$ of men and a set $W = \{w_1, w_2, \ldots, w_n\}$ of women. Each man ranks the women 1st through n^{th} via a 1-1 onto function $r_{m_i}\colon W \to \{1, 2, \ldots, n\}$, where m_i prefers w_j to w_k if $r_{m_i}(w_j) < r_{m_i}(w_k)$. Similarly, each woman ranks the men. The most basic problem is to find a *stable matching*; that is a 1-1 onto function $\mu\colon M \to W$ such that there is no man and woman each of whom prefers the other to his or her match assigned by μ.

The men-propose G-S algorithm produces a stable matching in finitely many rounds as follows: In each round each man who is not currently engaged proposes to his preferred woman among those who have not yet rejected him. Then each women who has been proposed to in this round tentatively accepts her preferred man among her current fiancée if she has one and those who have just proposed to her. The process ends when, at the end of a round, all women are engaged. Then all engagements become marriages.

The m-p G-S algorithm is men optimal and women worst in that for any preference profile pair there is no stable matching that provides any man with a more desired mate or any woman with a less desired mate.

3 Generating Intercorrelated Preferences.

Our method for generating men's and women's preferences with arbitrary levels of correlation and intercorrelation proceeds in two steps. In step one man $m_i \in M$ assigns a score $X_{i,j}$ to

woman $w_j \in W$ according to equation (1), which will be explained momentarily.

$$X_{i,j} = \eta_{i,j} + U_M I_j + V_M |i - j|_{100} \tag{1}$$

In step 2 each man m_i ranks the women first through last with $r_{m_i}(w_j) < r_{m_i}(w_k)$ if $X_{i,j} < X_{i,k}$. Then each woman w_j forms preferences over the men in a similar manner.

The first two terms on the right side of equation (1) follow the method of Caldarelli and Capocci (2001) for generating correlated preferences. Each of $\eta_{i,j}$ and I_j are randomly chosen from $[0, 1]$; $\eta_{i,j}$ represents the personal dimension of m_i's opinion of w_j; I_j represents the common opinion of w_j by all the men; U_M is arbitrarily chosen from $[0, +\infty)$ and provides a simple measure of the level of correlation of men's preferences, since increasing U_M tends to make the men's preferences more alike; $|i - j|_{100} = \min\{|i - j|, 100 - |i - j|\}/50$ can be interpreted as the distance from m_i to w_j if for each i m_i and w_i are placed at hour i on a 100-hour clock with circumference 2; and V_M is chosen arbitrarily from $(-\infty, +\infty)$. For large positive V_M, men tend to prefer closer women as in Example 1 of Section 1, while for large negative V_M, men tend to prefer distant women, as in Example 2. Therefore, for V_W replacing V_M when women's preferences are formed using equation (1) with suitable alterations, $V_M \times V_W$ provides a simple measure of intercorrelation–when V_M and V_W have the same sign, preferences tend to be positively intercorrelated as in Examples 1 and 2, and when V_M and V_W have opposite signs, preferences tend to be negatively intercorrelated as in Example 3.

We end this section with a caveat. Correlation and intercorrelation are strongly intertwined, so that there is no hope that $V_M \times V_W$ is a perfect measure of intercorrelation. For example, if U_M is much larger than V_M and V_W, then men's preferences will almost certainly be perfectly correlated, that is, identical. When men's preferences are identical, there will in fact be no intercorrelation regardless of the size of $V_M \times V_W$.

4   Gender Satisfaction and Intercorrelation of Preferences.

In order to make a preliminary examination of the effect of intercorrelation of preferences on the levels of men's and women's satisfaction with the outcome of the men-propose G-S algorithm, we fixed $n = 100$ and formed the 320 ordered 4-tuples (U_M, U_W, V_M, V_W) with $U_M = U_W \in \{0, .5, 1, 2, 3\}$ and $V_M, V_W \in \{-3, -2, -1, -.5, .5, 1, 2, 3\}$. For each of the ordered 4-tuples, for 100 trials we generated men's and women's preferences as described in Section 3 and ran the men-propose G-S algorithm. Figure 1 displays the results with $V_M \times V_W$ on the horizontal axis and MS, the sum of men's rankings of their assigned mates (averaged over trials), on the vertical axis. Keep in mind when examining the figures that higher MS means lower men's satisfaction, since an individual man is less satisfied when matched with his 17th ranked woman, for example, then when matched with his 12th ranked woman. In Figure 1 we see that this first step in the analysis has done little to clarify the relationship between intercorrelation of preferences and men's aggregate satisfaction.

Figure 2 displays the results with $V_M \times V_W$ on the horizontal axis and WS, the sum of women's rankings of their assigned mates, on the vertical axis. Here as expected satisfaction increases, that is WS decreases, with increasing intercorrelation, as measured by $V_M \times V_W$. We also ran a regression of women's satisfaction against intercorrelation:

$$WS = \beta_0 + \beta_1 U_M + \beta_2 U_W + \beta_3 \exp(-\exp(-\beta_4(V_M \times V_W - \beta_5))) \tag{2}$$

5

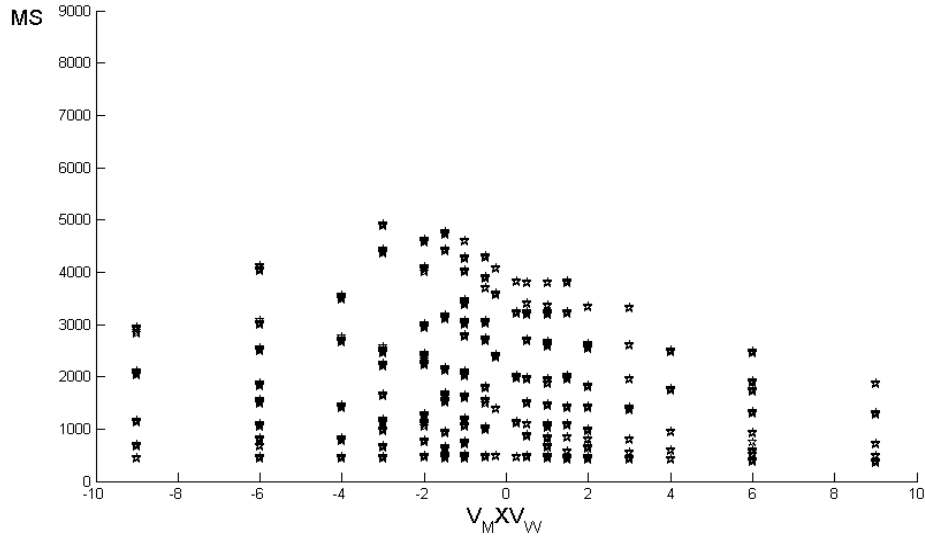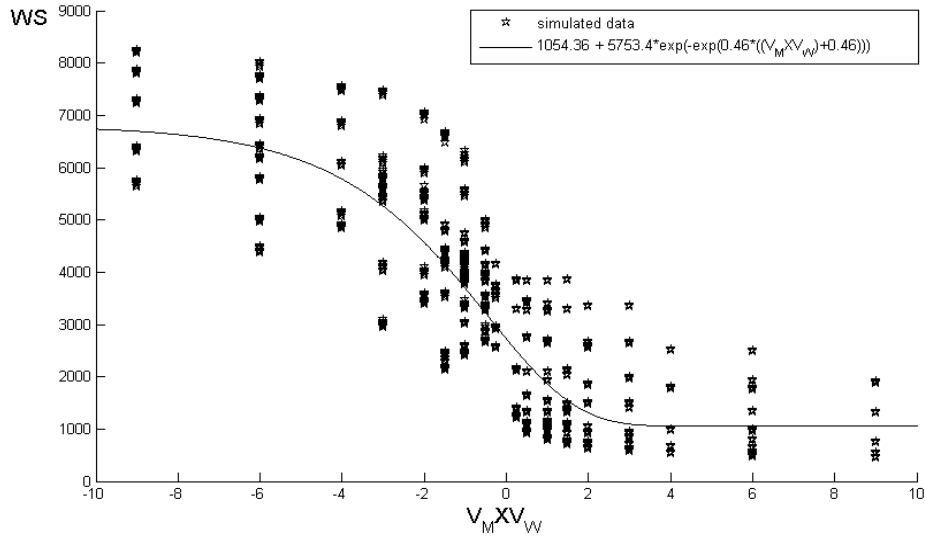Figure 1: Men's Satisfaction vs. $V_m \times V_w$

Figure 2: Women's Satisfaction vs. $V_m \times V_w$

Here the Gompertz function was chosen because a sigmoid function takes on the S shape that we expect of the data; we expect WS to increase with positive intercorrelation and to decrease with negative intercorrelation, and we know that WS is bounded below (by n) and above (by

Figure 3: Women's Satisfaction vs. $V_m \times V_w$, Separated by Correlation
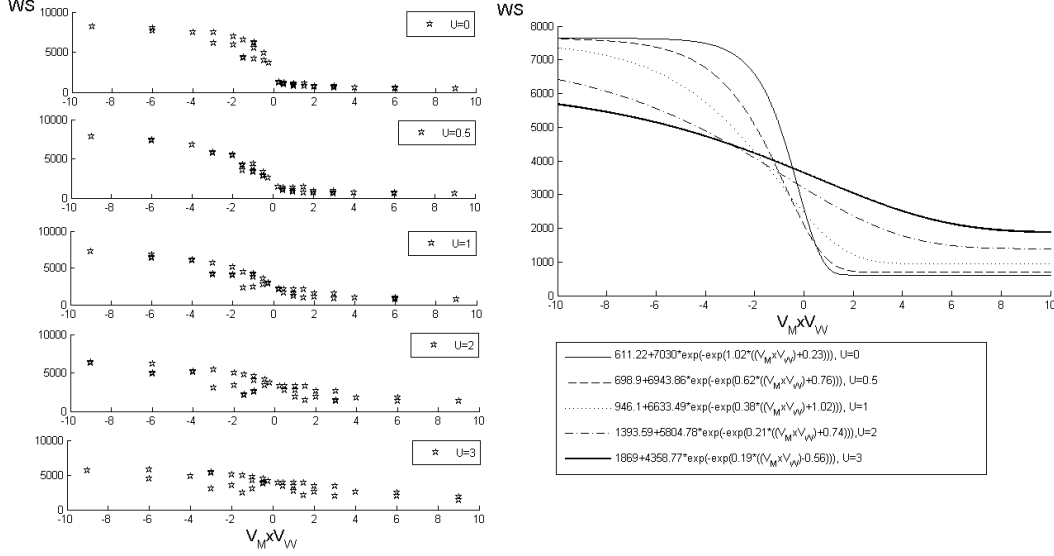
n^2). The corresponding estimation curve appears in Figure 2. The coefficients β_1 and β_2 are omitted in our figures because they are seldom significant. All of the coefficients presented are significant at the five percent level or better.

Two observations will guide us in refining our analysis. First, the scattering in Figures 1 and 2, that is, the vertical spread of the data on each vertical line containing data points, is in large part due to the fact that each such line contains points with correlations U_M and U_W that range from 0 to 3. We know from Celik and Knoblauch (2007) that correlation affects gender satisfaction.

Second, we had greater success with women than with men in the first step of our analysis, because there is a third factor that affects men's satisfaction. That third factor is V_M. To see why V_M is a greater factor in determining men's satisfaction than V_W is in determining women's satisfaction, consider the extreme case where V_M is so great in absolute value that no two men rank the same woman first. Since men propose, each man is matched with his first choice. In this extreme case men's satisfaction is entirely determined by V_M, while women's satisfaction depends on $V_M \times V_W$.

5   Isolating Intercorrelation.

In order to disentangle the effects of intercorrelation and correlation, we formed the 64 ordered 4-tuples (U_M, U_W, V_M, V_W) with $U_M = U_W = 0$ and $V_M, V_W \in \{-3, -2, -1, -.5, .5, 1, 2, 3\}$. For each ordered 4-tuple, for 100 trials we generated men's and women's preferences as described in Section 3 and ran the men-propose G-S algorithm.

We then ran regressions of women's satisfaction against intercorrelation using equation (2).

7

Figure 4: Men's Satisfaction vs. $V_m \times V_w$, Separated by Correlation
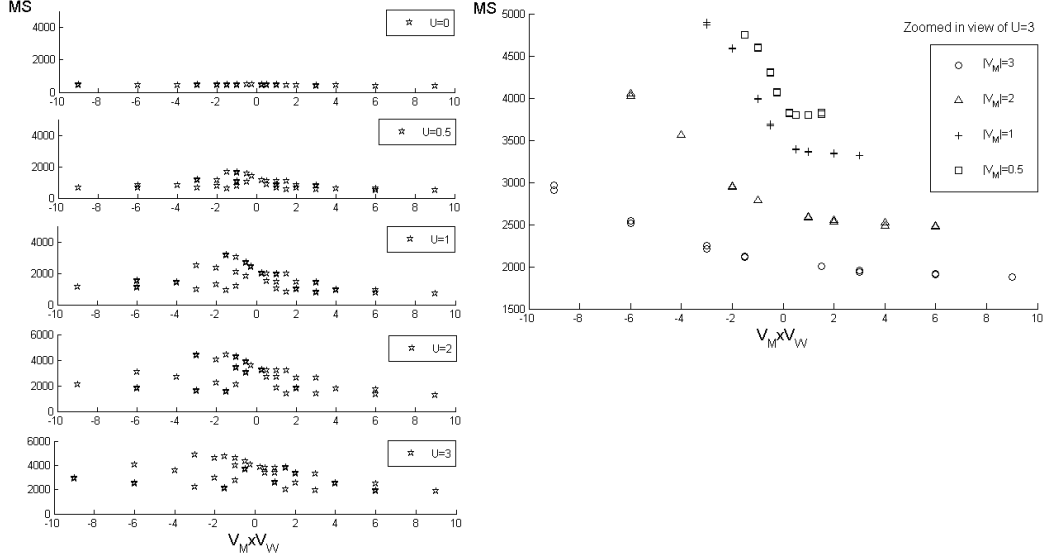
We repeated this procedure four more times with $U_M = U_W = U$ taking on successive values .5, 1, 2 and 3. The results are illustrated in Figure 3. On the left hand side of the figure are graphs of the data. It is clear that we have been successful in reducing the vertical scatter of the data, especially for lower values of U. On the right side of Figure 3 are the estimation curves for the five data sets on the left and a key that relates the curves to the data and supplies formulas for the curves. We can see in Figure 3 that correlation also alters the impact of intercorrelation:

1. For women the effect of intercorrelation decreases with increasing correlation, as can be deduced from the absolute variations of the curves in Figure 3.

2. For women the effective range of intercorrelation, that is, the range over which small changes in intercorrelation affect satisfaction, increases with increasing correlation, as can be seen from the curves.

Next we graphed the same data on the left side of Figure 4, but with MS on the vertical axes. Although using the five values of U to separate the data into five graphs has provided some reduction in scattering, it is clear that there must be some other factor besides correlation and intercorrelation of preferences that plays an important role in determining aggregate men's satisfaction. Guided by our observation at the end of Section 4, we further partitioned the data as follows: choosing the graph for $U = 3$ as an example, we enlarged that graph, placed it on the right side of Figure 4, and labeled the data points according to whether $|V_M| = .5, 1, 2$ or 3.

We conclude from Figure 4 that

1. correlation, intercorrelation and $|V_M|$, which we will call localization of preferences, together completely determine men's aggregate satisfaction with the m-p G-S matching, at least when preferences are generated using our method;

8

from the left side of Figure 4 that

2. men's satisfaction decreases with increasing correlation;

and from the right side of Figure 4 that

3. men's satisfaction increases with increasing localization of preferences,

4. men's satisfaction increases with increasing intercorrelation (this trend is difficult to see until the data has been partitioned by both U and $|V_M|$), and,

5. men's sensitivity to intercorrelation increases as $|V_M|$ decreases. Since decreasing $|V_M|$ with U held fixed tends to increase the true level of correlation, we can say men's sensitivity to intercorrelation increases as correlation increases. The opposite is true for women as we can see from the right side of Figure 3.

6   A General Measure of Intercorrelation.

The usefulness of $V_M \times V_W$ as a measure of intercorrelation of preferences is limited by the fact that V_M and V_W are only defined for preference profiles generated using the method of Section 3. We now define a measure of intercorrelation for any preference profile pair. First square the difference between the rank m_i gives w_j and the rank w_j gives m_i and add over all women.

$$\phi_{m_i} = \sum_{j=1}^{n}(r_{m_i}(w_j) - r_{w_j}(m_i))^2 \tag{3}$$

Then sum across the men and divide by n.

$$\phi_{ave} = \frac{\sum_{i=1}^{n}\phi_{m_i}}{n} \tag{4}$$

Now normalize using the maximum possible ϕ_{ave} score obtained when no two men agree on the rank of any woman, and each man is ranked last by his first-ranked woman, second last by his second ranked woman, third last by his third ranked woman, etc.

$$\Phi = \frac{\sum_{k=1}^{n}(n+1-2k)^2 - \phi_{ave}}{\sum_{k=1}^{n}(n+1-2k)^2} \tag{5}$$

Then perfect positive intercorrelation yields $\Phi = 1$ and perfect negative intercorrelation yields $\Phi = 0$.

The obvious question is, will Φ behave similarly to $V_M \times V_W$ as a determining factor of gender satisfaction with the outcome of the m-p G-S algorithm despite the one major difference between $V_M \times V_W$ and Φ? That difference is that Φ measures intercorrelation *after* preferences are formed, while $V_M \times V_W$ measures expected intercorrelation without taking into account what U_M and U_W are chosen. Then, for example, even if V_M and V_W are large, for sufficiently large U_M men's preferences will almost certainly be identical, so that Φ will almost certainly be $1/2$.

To answer this question we repeated the work of the previous section, except that we used Φ as a measure of intercorrelation rather than $V_M \times V_W$. Comparing Figure 5 with Figure 2, we see that Φ does a markedly better job of predicting WS than does $V_M \times V_W$. In Figure 6 we see that
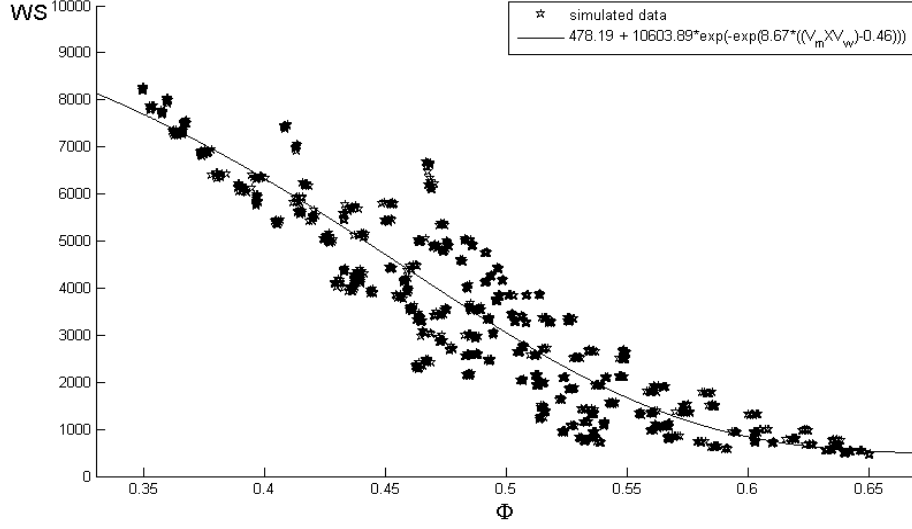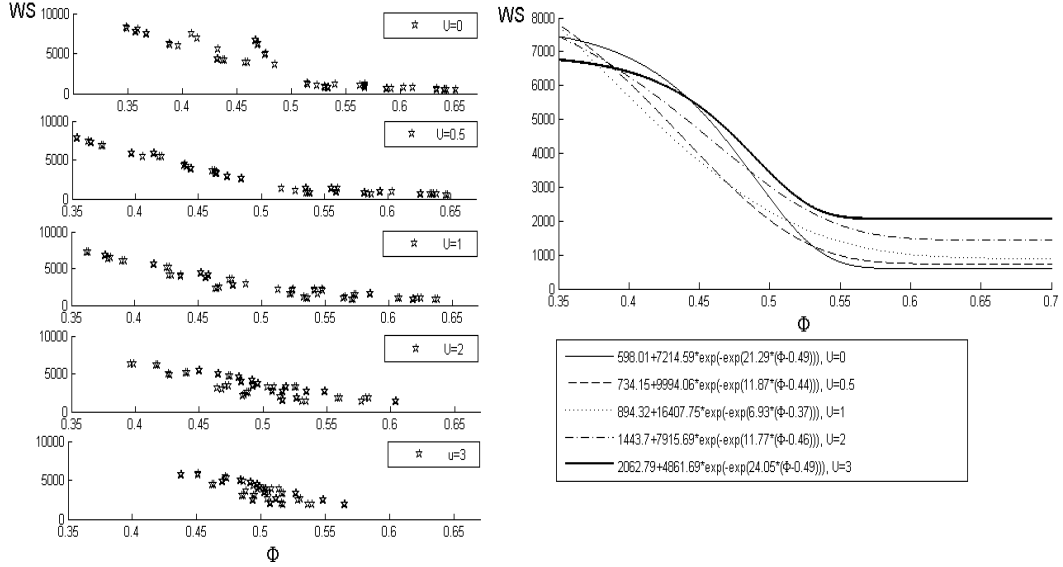
9

Figure 5: Women's Satisfaction vs. Φ

Figure 6: Women's Satisfaction vs. Φ, Separated by Correlation

the two observations in Section 5 concerning women do not hold up or are much less pronounced than indicated in Figure 3. Also in Figure 6 it can be seen that there is an asymmetric aspect to the impact of intercorrelation. Negative intercorrelation has a gradual effect on WS, but as

Figure 7: Size of Stable Set vs. Φ, Separated by Correlation
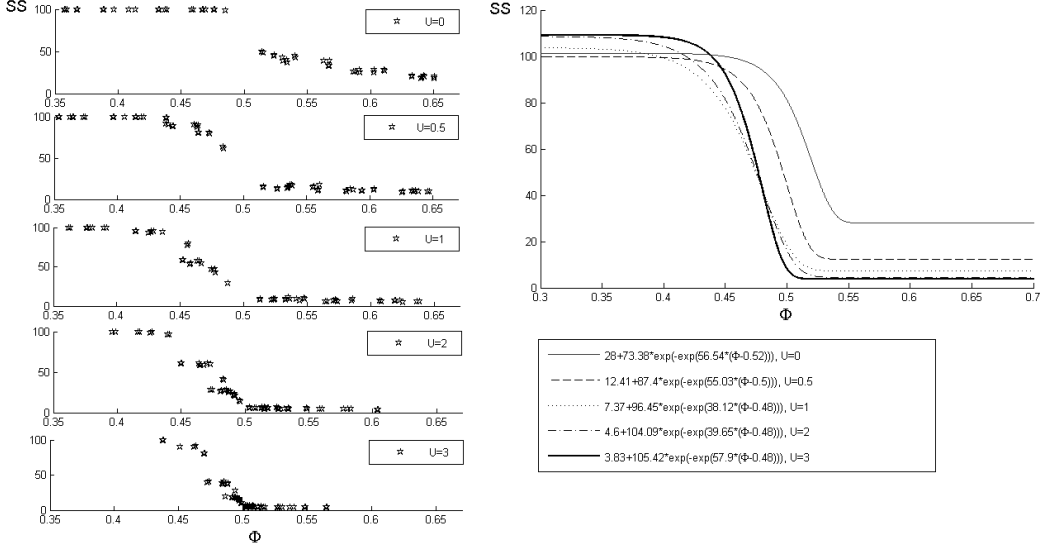
soon as intercorrelation is positive, the estimation curves level off quickly. We attribute these differences to the fact Φ is a measure of intercorrelation, while $V_M \times V_W$ is a measure of expected intercorrelation that does not take into account the suppression of intercorrelation by high values of U.

On the other hand Φ and $V_M \times V_W$ are very similar as predictors of aggregate *men's* satisfaction, and the five observations in Section 5 that concern men hold up under analysis via Φ. There is no reason to include the graphs that we compared with Figures 1 and 4 to reach these conclusions.

7 Strategic Reporting of Preferences.

In this section we investigate the effect of intercorrelated preferences on the susceptibility of the m-p G-S algorithm to strategic reporting of preferences. We follow Roth and Peranson (1999) in using the number of women who would match with a different partner if the women propose G-S algorithm replaced the m-p G-S algorithm as a proxy for susceptibility of the m-p G-S algorithm to strategic reporting of preferences. We refer to our proxy as SS for short, where SS stands for the size of the stable set.

The result of generating preferences and then matching via the m-p G-S algorithm is presented in Figure 7. We see that susceptibility to strategic reporting of preferences decreases with increasing intercorrelation, and, for positively intercorrelated preferences, decreases with increasing correlation.

Comparing Figure 6 and 7, we see that as was the case with aggregate women's satisfaction, intercorrelation and correlation together are good predictors of susceptibility to strategic report-

11

ing of preferences; and strategic reporting of preferences is even more sensitive to intercorrelation than is women's satisfaction, and this sensitivity occurs over a shorter interval of intercorrelation.

8 Concluding Remarks.

We have shown that correlation and intercorrelation of preferences and localization of proposers' preferences are the major determining factors of aggregate proposers' and proposees' satisfaction with the Gale-Shapley outcome.

We illustrate some of our findings by applying them in a discussion of the effects of legally mandated school preference for applicants who live nearby and/or have siblings already enrolled on the outcome of the student-propose Gale-Shapley admissions matching. As we mentioned in Example 1, such legal mandates introduce positive intercorrelation.

From our observations in Sections 4, 5 (observation 4) and 7, introduction of legal mandates will increase aggregate school and aggregate student satisfaction, and will decrease susceptibility to strategic reporting of preferences. By Section 5 observation 5, and the right side of Figure 4, the magnitude of the positive effect of the legal mandate on student satisfaction is dependent on correlation of preferences, localization of student preferences and degree of intercorrelation before and after legal mandates go into effect.

References

[1] Abdulkadíroğlu, A, P. Pathak & Roth, A.E. (2005) The New York City High School Match. *American Economic Review*, *95*(2), 364–367.

[2] Abdulkadíroğlu, A, P. Pathak, Roth, A.E. & T. Sönmez (2005) The Boston Public Schools Match. *American Economic Review*, *95*(2), 368–371.

[3] Celik, O. B., V. Knoblauch (2007) Marriage Matching with Correlated Preferences. *Working Paper*, University of Connecticut.

[4] Chen, Y., T. Sonmez (2006) School Choice: An Experimental Study. *Jounral of Economic Theory*, *127*, 202–231.

[5] Ehlers, L., J. Masso (2005) Incomplete Information and Small Cores in Matching Markets. *Working Paper*, Universite de Montreal, 1–25.

[6] Gale, D., L.S. Shapley (1962) College Admissions and the Stability of Marriage. *Amer. Math. Monthly*, *69*(1), 9–15.

[7] Haruvy, E., A.E. Roth, & Unver, M.U. (2006) The dynamics of law clerk matching: An experimental and computational investigation of proposals for reform of the market. *Journal of Economic Dynamics & Control*, *30*, 457–486.

[8] Knoblauch, V. (accepted 2008) Marriage matching and gender satisfaction. *Social Choice and Welfare*, *forthcoming* DOI 10.1007/s00355-008-0303-2.

[9] Mongell, S., A.E. Roth (1991) Sorority Rush as a Two-sided Matching Mechanism. *American Economic Review*, *81*(3), 441–464.

[10] Roth, A.E., E. Peranson (1999) The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design. *American Economic Review, 89*(4), 748–780.

[11] Roth, A.E., M. Sotomayor (1990) *Two-sided Matching: A Study in Game-Theoretic Modeling and Analysis.* Cambridge University Press, Cambridge, MA.

[12] Sonmez, T. (1999) Strategy-Proofness and Essentially Single-Valued Cores. *Econometrica, 67*(3), 677–689.

[13] Teo, C.P., J. Sethuraman, & Tan, W.P. (2001) Gale-Shapley Stable Marriage Problem Revisited: Strategic Issues and Applications. *Management Science, 47*(9), 1252–1267.

[14] Wilson, L.B. (1972) An analysis of the marriage matching assignment algorithm. *BIT, 12*, 569–575.