DISCUSSION PAPER SERIES

IZA DP No. 7681

Household Behavior and the Marriage Market

Daniela Del Boca
Christopher J. Flinn

October 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Household Behavior and the Marriage Market

Daniela Del Boca
Universita di Torino,
Collegio Carlo Alberto and IZA

Christopher J. Flinn
New York University
and Collegio Carlo Alberto

Discussion Paper No. 7681
October 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7681
October 2013

ABSTRACT

Household Behavior and the Marriage Market*

There is some controversy in the field of household economics regarding the efficiency of household decisions. We make the point that a flexible specification of spousal preferences and the household production technology precludes the possibility of using revealed preference data on household time allocations to determine the manner in which spouses interact: efficiently or inefficiently. Under strong, but standard, assumptions regarding marriage market equilibria, marital sorting patterns can be used essentially as "out of sample" information that allows us to assess whether household behavior is efficient or not. We develop a new likelihood-based metric to compare marriage market fits under the two alternative behavioral assumptions. We use a sample of households drawn from a recent wave of the Panel Study of Income Dynamics, and find strong evidence supporting the view that household behavior is (constrained) efficient.

JEL Classification: D13, J12, J22

Keywords: bilateral matching, household time allocation, efficient outcomes, likelihood analysis

Corresponding author:

Daniela Del Boca
Department of Economics
University of Turin
Via Po 53
Torino 10124
Italy
E-mail: dani.delboca@unito.it

* This paper was previously circulated under the title "Household Time Allocation and Modes of Behavior: A Theory of Sorts." This research was partially supported by the C.V. Starr Center for Applied Economics at New York University and Collegio Carlo Alberto. Flinn thanks the National Science Foundation for additional support. The editor, an anonymous referee, and an anonymous associate editor provided constructive comments and suggestions that led to a substantially improved paper. We are grateful to Olivier Bargain, Cristian Bartolucci, Xiaohong Chen, Olivier Donni, Bryan Graham, James Mabli, Tom MaCurdy, Al Roth, and Yoram Weiss for helpful discussions and suggestions, as well as to participants in the August 2005 SITE workshop on "The Nexus Between Household Economics and the Macroeconomy," the conference "Interactions within the Family" held at the Universita di Torino in 2005, the conference "The Labour Market Behaviour of Couples" held in Nice in 2008, and workshop participants at ASU, Yale, and UCL. We remain responsible for all errors and omissions.

1 Introduction

Most analyses of household behavior conducted at the microeconomic level posit cooperative behavior by spouses (for some exceptions, see Chen and Woolley (2001) and Del Boca and Flinn (2012)). In fact, Chiappori and his coauthors (e.g., Chiappori (1992), Browning and Chiappori (1998)) have argued that all such models should posit efficiency as an identifying assumption when attempting to estimate individualistic preferences using data on household allocations. Such an assumption, however, leads to other difficult identification issues since the dependent variables, which are household allocations, are not uniquely determined without further auxiliary assumptions regarding how the household selects one particular efficient allocation from the continuum of such choices that typically exist.

Chiappori and his collaborators (e.g., Chiappori (1988,1992), Browning et al. (1994), Browning and Chiappori (1998), Bourguignon et al. (2009)) have proposed a data-based strategy to estimate the household utility function $\alpha(z)U_1(x) + (1 - \alpha(z))U_2(x)$, where $\alpha(z)$ is the Pareto weight attached to the individualistic utility of agent 1, x is a vector of consumption choices, and z is a vector of personal, household, and environmental characteristics. The solution to this problem lies on the Pareto frontier for $\alpha(z) \in [0, 1]$. Model identification is achieved through restrictions regarding the arguments of the weighting and individualistic utility functions and/or functional form assumptions. Identification is achieved without resort to a specific axiomatic solution, with the data (z and x) given the power to solve the multiple equilibria problem within the particular model structure.

In Del Boca and Flinn (2012), we showed that when allowing unrestricted individual heterogeneity in wages, preferences, and household productivity, models of noncooperative and cooperative behavior were nonparametrically identified (i.e., from information on wages, nonlabor incomes, and time allocation decisions of households) and that they were all simply different mappings of the data into the parameter space.[1] In an empirical sense, then, all of these models were equivalent. We then constructed a model based on Folk Theorem results that allowed households to choose their mode of behavior (cooperative or noncooperative). We showed that this model was not nonparametrically identified, and proceeded to estimate all of the models under parametric assumptions on the distributions of preferences and productivities. The "endogenous household interaction" model, as we called the model in which households choose to act cooperatively or not, was found to fit the data the best. The estimated parameters indicated that one-fourth of households behave in a noncooperative way with the rest using a cooperative decision rule.

In this paper we further explore the issue of the "mode" of household behavior, and for simplicity focus on only two alternatives, noncooperative Nash equilibrium (NE) and "constrained" Pareto optimal (CPO) behavior, to be defined below. We first show, as in our earlier paper, that after allowing for general forms of population heterogeneity in

[1] The models are all saturated models in which the number of parameters (fixed preference and productivity draws) is equal to the number of data points. They amount to different parameterizations of a saturated model.

preferences and household productive ability, it is not possible to distinguish between NE and CPO solely on the basis of household time allocation data. To do so requires imposing homogeneity restrictions that may not be justifiable and are essentially untestable. We then show how patterns of *marital sorting* observed in the data potentially contain valuable information on the manner in which household members interact. We are by no means the first to point this out. Following the view of Becker (1991) that marriage is a partnership for joint production and consumption, several authors have analyzed aspects of the marriage market to explore marital behavior and the gains to marriage (e.g., Choo and Siow (2006), Dagsvik et al. (2001), Pollack (1990)). Other research has explored the effects of the marriage market on household behavior. While Aiyagari, Greenwood and Guner (2000) and Greenwood, Guner and Knowles (2003) have focused on the link between the marriage market and parental investments in children and patterns of intergenerational mobility, Fernandez et al. (2005) have studied the implication of marital sorting for household income inequality.

Microeconomic analyses such as Browning et al. (2003), Seitz (1999), and Igiyun and Walsh (2007) have explored aspects of household formation that precede marriage to merge household models with marital sorting in order to explore the implications of spousal matching for intrahousehold allocations. While the objective of these papers is mainly to identify sharing rules and to consider whether household allocations are efficient, we use marital sorting to investigate *which* type of interaction is most consistent with observed outcomes.[2]

The basic idea of our approach can be summarized as follows. We begin by assuming that spouses make household allocation decisions using some rule R, and then use the observed household time allocations, along with the observed wages and nonlabor incomes, to "back out" the parameters characterizing the preferences and (household) productivities of both spouses within each household in the sample. Using these individual-specific parameters, we can then construct preference orderings for each male over all possible females in a particular marriage market assuming that the household allocations are chosen according to R, and we can construct the preference orderings for the females in a similar manner. Armed with these R-specific preference orderings, we then apply the Gale and Shapley (1962) - henceforth GS - bilateral matching algorithm to determine the predicted stable pairings under R. We compare the correspondence between the predicted matches and the observed ones for R using a likelihood-based metric. The likelihood function we propose is defined over a parameter space of household allocation rules, $\Re$. On a conceptual level, the likelihood function defines a metric over which we can compare any potential allocation rules $R, R' \in \Re$, and, in principle, allows us to choose a rule which maximizes the likelihood

[2] A paper by Iyigun (2007) considers the choices of couples of whether to implement efficient outcomes within a marriage market context. His objective is to examine whether there exist equilibria in which some households behave efficient while others do not. This paper uses household behavior and marriage market outcomes to infer whether all households are behaving efficiently or all households are not. Thus our framework is less general in that sense, but this is due to the empirical focus of this paper.

of the spousal matches observed in the data. Due to identification problems associated with the first stage of the estimation problem, that in which the preferences and productivities of spouses are determined, we limit our empirical work to a comparison of the allocation rules associated with NE and CPO.

Before we begin, a few words concerning some of the modeling choices that have been made. Since the allocation rules utilized within households have been derived within a static framework, we chose only to consider a static marriage market equilibrium as well. The most appealing alternative to us was a dynamic model of the search for partners. While the most basic search framework is capable of producing stationary equilibrium distributions of match types that are uniquely determined (and hence avoids the possibility of multiple equilibria which typically exist using the GS algorithm), this comes only under common yet extreme assumptions on the matching technology (e.g., non-directed search). Since the relationship between household outcomes and sorting in the marriage market is relatively immediate under a static behavioral model and the GS matching algorithm, we opted for this pairing. It is by no means the only choice that could have been made.

In terms of functional form assumptions, we have utilized Cobb-Douglas utility functions for the spouses and a Cobb-Douglas production technology for producing the household (public) good. The simple forms of the first order conditions associated with these functions make the econometric identification arguments relatively straightforward. In terms of links to the empirical literature on household allocation decisions, Cobb-Douglas functional form assumptions are among the most commonly utilized (in addition to the more general CES or CRRA).

Using a sample of married couples drawn from the Panel Study of Income Dynamics (PSID) in 2007, we conduct a test between the cooperative and noncooperative models based on what is essentially a likelihood ratio statistic. The test is repeated for regional marriage markets and for different sizes of marriage choice sets, which are defined below. In virtually all of the tests, the cooperative model (CPO) outperforms the noncooperative one (NE) in the sense of correctly predicting marital sorting patterns. Thus the results of this analysis are largely consistent with those of Del Boca and Flinn (2012), even though the tests are based on very different features of the data.

The plan of the paper is as follows. Section 2 contains the description of the model and the bilateral matching algorithm. In Section 3 we explore econometric issues, including the nonparametric estimation of the distribution of state variables and the formation of a likelihood function for the spousal pairings observed in the data given the first stage estimates of the distribution of state variables. Empirical results are presented in Section 4, and Section 5 contains a brief conclusion.

2 Model

The focus of our attention will be household formation. Without loss of *empirical* generality (as we shall see below), we will assume that spouse i has the following simple determination of his or her utility in a static context. We assume a Cobb-Douglas utility function for individual i of the form

$$u_i(l_i, K) = \lambda_i \ln(l_i) + (1 - \lambda_i) \ln(K),\ i = 1, 2,$$

where l_i is the leisure of individual i, K is a public good that is produced within the household, and λ_i is the preference weight attached to leisure, which is the only private consumption good in the model. The household good K is produced according to the constant returns to scale Cobb-Douglas technology

$$K = \tau_1^{\delta_1} \tau_2^{\delta_2} M^{1-\delta_1-\delta_2},$$

where τ_i is the time input of spouse i in household production, δ_i is the elasticity of K with respect to time input τ_i, and M is total income of the household, or

$$M = w_1 h_1 + w_2 h_2 + y_1 + y_2,$$

where w_i is the wage rate of spouse i, h_i is their hours of work, and y_i is the nonlabor income of spouse i. The "physical" time endowment of each spouse is T, and

$$T = l_i + h_i + \tau_i,\ i = 1, 2.$$

Each individual has their own value of market productivity, with the value of their time in the market given by w_i. Moreover, each individual has a nonlabor income level of y_i. Both of these quantities are determined outside of the model.

The population is characterized by heterogeneity in all of the parameters that appear in the functions defined above. The population consists of two types of agents, males (potential husbands) and females (potential wives). Each subpopulation (defined with respect to gender) is characterized by a distribution of characteristics particular to that type. The cumulative distribution function of characteristics of individuals of gender i is

$$G_i(\lambda_i, \delta_i, w_i, y_i).$$

Then a household is defined by the vector of state variables

$$S = (\lambda_1, \delta_1, w_1, y_1) \cup (\lambda_2, \delta_2, w_2, y_2).$$

Given a value of S, the household determines equilibrium time allocations and the resultant welfare distribution in the household according to some rule R. Thus R is a mapping from S into a vector of observable household choices, in our case given by the vector

$$C = (h_1, h_2, \tau_1, \tau_2).$$

Thus

$$C = R(S). \tag{1}$$

We will discuss specific properties of the mapping R below, but for now we assume that R assigns a unique value C to any vector $S \in \Omega_S$, where we will think of Ω_S as the space of household characteristics.

2.1 Noncooperative Behavior

We begin our investigation of the time allocation decision of the household with the case of Nash equilibrium. Later we will turn our attention to cooperative models of household behavior.

The reaction function for spouse 1 in a household characterized by S is given by

$$\begin{aligned}(h_1^*, \tau_1^*)(h_2, \tau_2; S) &= \arg\max_{h_1, \tau_1} \lambda_1 \ln(T - h_1 - \tau_1) \\ &\quad + (1-\lambda_1)[\delta_1 \ln \tau_1 + \delta_2 \ln \tau_2 + (1 - \delta_1 - \delta_2) \ln(y + w_1 h_1 + w_2 h_2)],\end{aligned}$$

where $y = y_1 + y_2$. Assuming an interior solution for h,[3] the solutions are given by continuously differentiable functions

$$\begin{aligned} h_1^* &= h_1^*(h_2, \tau_2; S) \\ \tau_1^* &= \tau_1^*(h_2, \tau_2; S). \end{aligned}$$

An analogous pair of reaction functions exists for the second individual. Under our specification of preferences and the production technology, there exists a unique Nash equilibrium

$$\begin{aligned} h_1^{**} &= h_1^*(h_2^{**}; \tau_2^{**}; S) \\ \tau_1^{**} &= \tau_1^*(h_2^{**}, \tau_2^{**}; S) \\ h_2^{**} &= h_2^*(h_1^{**}, \tau_1^{**}; S) \\ \tau_2^{**} &= \tau_2^*(h_1^{**}, \tau_1^{**}; S). \end{aligned}$$

Insuring that h_1^{**} and h_2^{**} are both greater than zero requires restricting the parameter space Ω_S. We will provide further discussion of this point in the following section.

Associated with the Nash equilibrium is a welfare pair $(V_1^N, V_2^N)(S)$. These values will be used as outside options in the constrained Pareto efficient allocation we discuss next. After considering the marital sorting process, we will justify the use of these values as forming a participation constraint for any efficient allocation.

[3] Whenever $\alpha_1 > 0$ and $\delta_1 > 0$, an interior solution for τ_1 is assured by the Inada condition.

2.1.1 Pareto Optimal Decisions

We write the Benthamite social welfare function for the household as

$$W_\alpha(a_1, a_2) = \alpha u_1(a_1, a_2) + (1 - \alpha) u_2(a_1, a_2),$$

where $\alpha \in (0, 1)$ is the "notional" Pareto weight associated with spouse 1. Given the state vector S, which contains elements that describe the individual preferences and constraint set of the household, maximizing the value of $W_\alpha(a_1, a_2)$ with respect to the actions of the two spouses as we vary α from 0 to 1 traces out the Pareto frontier. In particular,

$$(\hat{a}_1^P \ \hat{a}_2^P)(\alpha; S) = \arg \max_{(a_1, a_2)} W_\alpha(a_1, a_2).$$

Assumption 2 The Pareto optimal actions $(\hat{a}_1^P \ \hat{a}_2^P)(\alpha; S)$ are unique.

The Pareto frontier is defined by the set of points

$$\begin{gathered} PF(S) \equiv \\ \{u_1(\hat{a}_1^P(\alpha; S), \hat{a}_2^P(\alpha; S)), u_2(\hat{a}_1^P(\alpha; S), \hat{a}_2^P(\alpha; S))\}, \\ \alpha \in [0, 1] \end{gathered}$$

It follows that u_1 is nondecreasing in α and u_2 is nonincreasing in α along PF. Assuming differentiability along the Pareto frontier with respect to α,

$$\left. \frac{du_2}{du_1} \right|_{\{u_2, u_1\} \in PF} < 0.$$

Our analysis is based on the existence of externalities within the household. It follows that the pair of Nash equilibrium utility levels, $\{V_1^N, V_2^N\}$, is not a point on the Pareto frontier.

2.1.2 Constrained Pareto Outcomes

Pareto efficient outcomes have the desirable feature that one spouse's utility cannot be improved without decreasing the other's, but may or may not meet certain reasonable fairness criteria. We singled out Nash equilibrium play as a natural focal point for the behavior of spouses since it involves no coordination or monitoring of actions due to the fact that strategies are best responses, at least in static games. This gives the Nash equilibrium a type of stability not possessed by efficient outcomes, at least in a problem for which Folk Theorem results cannot be brought to bear.

At a minimum, then, it seems reasonable to restrict attention to efficient outcomes that yield each spouse at least the level of welfare they can attain under the Nash equilibrium actions. We think of this as a "short-run" participation constraint, where short-run refers to the fact that it ensures that each party is at least as well-off under the efficient allocation as they would be under Nash equilibrium in the current period.[4,5] The following proposition follows directly from the relatively weak assumptions made to this point. To reduce notational clutter, we drop the explicit conditioning on the state vector S.

Proposition 1 *There exists a nonempty interval* $I^C(V_1^N, V_2^N) \equiv [\underline{\alpha}(V_1^N), \overline{\alpha}(V_2^N)] \subset (0,1)$, *with* $\underline{\alpha}(V_1^N) < \overline{\alpha}(V_2^N)$, *such that*

$$\begin{aligned} u_1(\hat{a}_1^P(\alpha), \hat{a}_2^P(\alpha)) &\geq V_1^N \\ u_2(\hat{a}_1^P(\alpha), \hat{a}_2^P(\alpha)) &\geq V_2^N \end{aligned}$$

if and only if $\alpha \in I^C(V_1^N, V_2^N)$.

Proof. The points on the Pareto frontier are monotone and continuous functions of α. By Assumption 2, the Nash equilibrium payoffs are dominated by a set of points on the Pareto frontier. Define the unique value $\underline{\alpha}(V_1^N)$ by

$$u_1(\hat{a}_1^P(\underline{\alpha}(V_1^N), \hat{a}_2^P(\underline{\alpha}(V_1^N))) = V_1^N,$$

and, similarly, define $\overline{\alpha}(V^N)$ by

$$u_2(\hat{a}_1^P(\overline{\alpha}(V_2^N)), \hat{a}_2^P(\overline{\alpha}(V_2^N))) = V_2^N.$$

Define the set of all α values that yield an efficient payoff at least as large as V_1^N for spouse 1 as

$$I_1^C(V_1^N),$$

the minimum element of which is $\underline{\alpha}(V_1^N)$. Define the set of all values of α that yield an efficient payoff at least as large as V_2^N to spouse 2 as

$$I_2^C(V_2^N),$$

the maximum element of which is $\overline{\alpha}(V_2^N)$. Then

$$\begin{aligned} I^C(V_1^N, V_2^N) &= I_1^C(V_1^N) \cap I_2^C(V_2^N) \\ &= [\underline{\alpha}(V_1^N), \overline{\alpha}(V_2^N)] \\ &\neq \varnothing \end{aligned}$$

[4] In Del Boca and Flinn (2012), we also consider "long-run" participation constraints, which are those that satisfy the implementation conditions associated with Folk Theorem arguments.

[5] In his dynamic contracting view of household behavior, Mazzocco (2007) investigates the nature of the period by period participation constraints of the household. Under commitment, the household may commit to supplying each member with an average amount of utility over time rather than within each period. In such a case, the type of within period constraint we propose may not hold. However, in our static model of household allocation and marital sorting, this distinction does not arise.

■

Given problems associated with the identification of the welfare weight α, which are discussed in detail below and in Del Boca and Flinn (2012), we will assume that there exists one "notional" value of α, common to all marriages, which could be thought of as being culturally determined. The constrained Pareto optimal allocation is obtained by first determining whether $\alpha \in I^C(V_1^N, V_2^N)$. If so, each spouse's utility level using the Pareto weight of α exceeds their static Nash equilibrium utility level, and the constraint is not binding. Instead, if $\alpha < \underline{\alpha}(V_1^N)$, the Pareto efficient solution yields less utility to spouse 1 than the Nash equilibrium solution. To get this spouse to participate in the Pareto efficient solution, it is necessary to provide them with at least as much utility as they would obtain in the static Nash equilibrium, which means adjusting the Pareto weight up to the value $\underline{\alpha}(V_1^N)$. Conversely, if $\alpha > \overline{\alpha}(V_2^N)$, then the Pareto weight has to be adjusted downward to $\overline{\alpha}(V_2^N)$ to provide the incentive for the second spouse to participate in the household efficient allocation. The formal statement of the actions in this environment is as follows. If $\alpha \in I^C$, then

$$(\hat{a}_1^C\ \hat{a}_2^C)(\alpha) = (\hat{a}_1^P\ \hat{a}_2^P)(\alpha),\ \alpha \in I^C, \tag{2}$$

since the "participation constraint" is not binding. If $\alpha < \underline{\alpha}(V_1^N)$, so that spouse 1 would have a higher payoff in Nash equilibrium, the α must be "adjusted" up so that he has the same welfare in either regime. In this case,

$$(\hat{a}_1^C\ \hat{a}_2^C)(\alpha) = (\hat{a}_1^P\ \hat{a}_2^P)(\underline{\alpha}(V_1^N)),\ \alpha < \underline{\alpha}(V_1^N). \tag{3}$$

Conversely, if spouse 2 suffers utility-wise in the efficient allocation associated with α, the α must be adjusted downward, and we have

$$(\hat{a}_1^C\ \hat{a}_2^C)(\alpha) = (\hat{a}_1^P\ \hat{a}_2^P)(\overline{\alpha}(V_2^N)),\ \alpha > \overline{\alpha}(V_2^N). \tag{4}$$

Note that under this behavioral rule, there is still, in general, a continuum of possible solutions, associated with all values of α belonging to I^C.

2.2 Marital Sorting

The gender-specific distributions of state variables, G_1 and G_2, are assumed to exogenously determined. In our analysis we only consider the case of balanced populations of males and females (i.e., $N_1 = N_2$) and we assume that marriage dominates remaining single for all males and females.[6] Then let G_1 and G_2 denote the marginal distributions of males and females, respectively. Male i is defined by his vector of characteristics

$$m_i = (\lambda_{1i}, \delta_{1i}, w_{1i}, y_{1i}),$$

[6] To construct the welfare of an individual in the single state, we set the productivity of his or her "missing" spouse to 0, so that the production function $K = \tau_j^{\delta_j}(w_j h_j + Y_j)^{1-\delta_j}$ for single individual j. In the empirical application we find that the value of being married to any individual of the opposite sex dominates the value of being single for any agent in the sample. In the sequel we seldom discuss the option of remaining single since it is not relevant empirically.

while female j is defined by her characteristics vector

$$f_j = (\lambda_{2j}, \delta_{2j}, w_{2j}, y_{2j}).$$

Following GS, we consider the simple case in which their exists a marriage market in which individuals from the different subpopulations are matched one-to-one, all individual characteristics are perfectly observable, and the market clears instantaneously. Each male has preferences over possible mates, with the preference ordering of male m_i given by $P(m_i)$.

Similarly, the preference ordering of woman j is given by $P(f_j)$. In each case, the preference ordering amounts to a sequence of potential mates ranked in descending order, and may include ties. In addition, remaining single may dominate being married to certain individuals of the opposite sex, though in our application we will assume that the value of marriage to any partner of the opposite sex exceeds the value of remaining single (see the discussion in footnote 6). We will assume that this is also the case in this example. When $N = 5$, the preference ordering of male 4 may be given by

$$P(m_4) = f_3, f_1, f_2, f_5, f_4.$$

That is, male 4's first choice as a mate is female 3, followed by 1, 2, 5, and 4. The preferences of female 2 might be represented by

$$P(f_2) = m_4, m_1, m_3, m_2, m_5.$$

In this case, she prefers male 4, followed by males 1, 3, 2, and 5.

A marriage market is defined by $(M, F; P)$, where

$$P = \{P(m_1), ..., P(m_N); P(f_1), ...P(f_N)\}$$

is the collection of preferences in the population, $M = \{m_1, ..., m_N\}$, and $F = \{f_1, ..., f_N\}$. Then we have the following:

Definition 2 *A matching μ is a one-to-one correspondence from the set $M \cup F$ onto itself of order 2 (that is $\mu^2(x) = x$) such that $\mu(m) \in F$ and $\mu(f) \in M$. We refer to $\mu(x)$ as the mate of x.*

The notation $\mu^2(x) = x$ is read as $\mu(\mu(x))$, and just means that the mate of individual $x's$ mate is individual x.

Definition 3 *The matching μ is individually rational if each agent is acceptable to his or her mate. That is, a matching is individually rational if it is not blocked by any (individual) agent.*

This is a weak concept, for it only requires that any spouse is better than the outside option of remaining single, which is the case in our application. A stronger notion is that of stability. Say that a matching μ has resulted in $\mu(m_i) = f_j$ and $\mu(f_k) = m_l$, but that male i strictly prefers f_k to f_j and female f_k strictly prefers m_i to m_l. Then the pair (m_i, f_k) can deviate from the matching assignment μ and improve their welfare. Such a match is unstable in the terminology of GS.

Definition 4 *A matching μ is stable if it is not blocked by any individual or any pair or agents.*

The main achievement of GS was to set out an algorithm for finding an equilibrium of the marriage game that was decentralized and constructive in the sense of establishing that at least one stable matching equilibrium exists. They assumed that preferences of agents was public information and the existence of a convention regarding the meeting and offering technology. Roth and Sotomayer (1990) devote considerable attention to the design of mechanisms that elicit truthful revelation of preference orderings when preferences are not public information, and also explore alternative meeting and proposal technologies. These important issues will be of less importance to us here given the nature of the application and the econometric and empirical focus of our analysis.

In our application a male individual i is characterized by the vector m_i. His induced preference ordering over the females $f_1, ...f_N$ is determined by R in the following manner. If m_i and f_j are matched, then the household is characterized by

$$S_{ij} = m_i \cup f_j.$$

Then equilibrium time allocations in the household are given by

$$C_{ij}(R) = R(S_{ij}).$$

Given our assumptions regarding the form of the "payoff" functions to i and j, we can define the value to m_i of being matched with f_j under R as

$$\begin{aligned} V_i(j; R) &= \lambda_{1i} \ln(l_1^*(S_{ij}; R)) + (1 - \lambda_{1i})\{\delta_{1i} \ln \tau_1^*(S_{ij}; R) + \delta_{2j} \ln \tau_2^*(S_{ij}; R) \\ &\quad +(1 - \delta_{1j} - \delta_{2j}) \ln(w_{1i} h_1^*(S_{ij}; R) + w_{2j} h_2^*(S_{ij}; R) + y_{1i} + y_{2j})\}. \end{aligned}$$

Given behavioral mode R, the preference ordering of i is given by

$$P(m_i|R) = f^i_{(1)}(R), f^i_{(2)}(R), ..., f^i_{(N)}(R),$$

where

$$V_i(f^i_{(1)}(R); R) > V_i(f^i_{(2)}(R); R) > ... > V_i(f^i_{(N)}(R); R).$$

the preference orderings of all of the females in the marriage market are determined in an analogous manner. Given knowledge of m_i, f_j, and R, the preference ordering of all population members is determined. This implies the following.

Definition 5 *A marriage market is defined by* $(M, F; R)$.

An equilibrium assignment is a function of marriage market characteristics. Then the set of stable matchings is determined by the characteristics vectors M and F and the behavioral model R, or $\Theta(M, F; R)$. Now there may exist, and generally do exist, multiple stable assignment equilibria. Among this set of equilibria, attention has focused on the two "extreme" stable matchings, the one that is most beneficial to men and the one most beneficial to women.[7] The GS matching algorithm, which they termed "deferred acceptance," enables one to determine at least these two, of the many possible, equilibria in a straightforward manner. We describe the computation of the male-preferred equilibrium. In a given round,

1. Each male not tentatively matched with a female makes a marriage proposal to the woman he most prefers among the set of women who have not rejected a previous proposal of his. If he prefers the state of being single to any of the women in his choice set, he makes no offer.

2. Each woman (tentatively) accepts the proposal that yields the maximum payoff to her from the set of offers made to her during the round plus and value of the match she carries over from the previous round, if she has one. Any man whose offer is refused in the period cannot make another marriage proposal to the woman rejecting him in future rounds.

3. The process is repeated until no man makes a new marriage proposal to any woman.

The female preferred stable matching equilibrium is found in the identical way after reversing the roles of two sexes as proposers and responders.

There may well exist other stable matchings besides these two. Given the generality of the preference structure, the size of the individual characteristic space, and the number of individuals in the marriage market in our empirical analysis, it is not possible to attempt to enumerate all possible stable matchings. We have computed the predicted marriage assignments using estimates of the state vectors m_i and f_j under the two R that we consider. We found that the same pairs were matched in over 98 percent of the cases in the male-preferred and female-preferred matchings. As a result, we use pairings from the male-preferred equilibria only in all of the empirical work that follows. The reader should bear in mind that other equilibria exist, even if they are not so different in the likelihood-based metric we use to assess the ability of the models to fit the actual marriage sorting patterns that we observe.

[7] When there is a unique equilibrium these stable matchings are identical, of course. A sufficient condition for this to be the case is if the preference orderings of all males and all females are identical.

2.3 Adding Match Heterogeneity

For econometric purposes, and also so as to more generally model marriage market behavior, we extend the previous formulation of the payoffs associated with a given couple marrying to include random match-specific heterogeneity. In particular, we write the utility of male i in a match with female j as

$$\tilde{u}_i(a_i, a_j, \varepsilon_{ij}) = u_i(a_i, a_j) + \varepsilon_{ij},$$

while the utility for female j is given by

$$\tilde{u}_j(a_i, a_j, \varepsilon_{ij}) = u_j(a_i, a_j) + \varepsilon_{ij},$$

where ε_{ij} is considered the "psychic" component of the match valuation to each member of the couple, and the state variables S_{ij} have been omitted for notational simplicity. We assume that the random match value for potential couple (i, j) is independently and identically distributed according to the absolutely continuous distribution F. The following result is relatively obvious, but given its importance for the econometric procedures we implement it is stated formally.

Proposition 6 *The valuation of a potential match between male i and female j under behavioral rule R, $R \in \{NE, CPO\}$, is given by*

$$\begin{aligned} \tilde{V}_i^R(\varepsilon_{ij}) &= V_i^R + \varepsilon_{ij} \\ \tilde{V}_j^R(\varepsilon_{ij}) &= V_j^R + \varepsilon_{ij} \end{aligned} \tag{5}$$

Proof. In terms of the reaction functions that define the Nash equilibrium outcome, since

$$\begin{aligned} \hat{a}_i(a_j) &= \arg\max_{a_i|a_j} \tilde{u}_i(a_i, a_j) = \arg\max_{a_i|a_j} u_i(a_i, a_j) \\ \hat{a}_j(a_i) &= \arg\max_{a_j|a_i} \tilde{u}_j(a_i, a_j) = \arg\max_{a_j|a_i} u_j(a_i, a_j), \end{aligned}$$

then the reaction functions are invariant with respect to ε_{ij}, and thus the Nash equilibrium actions are invariant with respect to the ε_{ij} draw. It follows that the total valuation of a match between i and j under Nash equilibrium is given by

$$\{\tilde{V}_i^N, \tilde{V}_j^N\} = \{V_i^N + \varepsilon_{ij}, V_j^N + \varepsilon_{ij}\}.$$

For the CPO case, first note that the side constraint that any permissible surplus decision provide each spouse with at least the welfare associated with the Nash equilibrium allocation can be written

$$\begin{aligned} u_i(\hat{a}_i^P(\alpha), \hat{a}_j^P(\alpha)) + \varepsilon_{ij} &\geq \tilde{V}_i^N \\ \Rightarrow u_i(\hat{a}_i^P(\alpha), \hat{a}_j(\alpha)) + \varepsilon_{ij} &\geq V_i^N + \varepsilon_{ij} \\ \Rightarrow u_i(\hat{a}_i^P(\alpha), \hat{a}_j^P(\alpha)) &\geq V_i^N, \end{aligned}$$

for male i, which is the same constraint as in the case of $\varepsilon_{ij} = 0 \ \forall(i,j)$. The constraint for the welfare of female j is written symmetrically. The solution to the constrained Pareto optimal problem can be written as

$$\begin{aligned}(\hat{a}_i^P(\alpha), \hat{a}_j^P(\alpha)) &= \arg\max_{a_i,a_j} \alpha u_i(a_i, a_j) + (1-\alpha)u_j(a_i, a_j) + \alpha\varepsilon_{ij} + (1-\alpha)\varepsilon_{ij} \\ &= \arg\max_{a_i,a_j} \alpha u_i(a_i, a_j) + (1-\alpha)u_j(a_i, a_j)\end{aligned}$$

for all α that satisfy the participation constraint. Since the actions chosen are invariant to ε_{ij}, then so is $u_k(\hat{a}_i^P(\alpha), \hat{a}_j^P(\alpha))$, $k = i, j$, so that the participation constraints are invariant with respect to ε_{ij}. Then the optimization problem is invariant with respect to ε_{ij}, and

$$\{\tilde{V}_i^C, \tilde{V}_j^C\} = \{V_i^C + \varepsilon_{ij}, V_j^C + \varepsilon_{ij}\}.$$

■

While the value of ε_{ij} will have no impact on the actions taken by the spouses under either NE or CPO, marital sorting is not invariant with respect to the match-specific draws of ε. The presence of i.i.d. match heterogeneity enables us to provide a more continuous measure of the ability to predict observed marriage patterns for the two behavioral rules we consider, in addition to representing a somewhat more romantic view of the marriage market.

2.4 An Example

To fix ideas and provide some motivation for the econometric methods that follow, we consider the following simple example. Let the marriage market consist of three males and three females, who have the characteristics:

Characteristics

	State Variables			
Individual	λ	δ	w	Y
f_1	0.20	0.30	10.00	140.00
f_2	0.05	0.25	9.00	100.00
f_3	0.10	0.60	10.00	50.00
m_1	0.30	0.15	18.00	20.00
m_2	0.15	0.10	12.00	150.00
m_3	0.40	0.30	20.00	150.00

.

Assuming that marriage dominates the single state for all individuals, there are six possible marriage patterns. The baseline payoffs, that is, the deterministic portion of the match values, for the females are:

Female Payoffs

	NE			CPO		
Male	f_1	f_2	f_3	f_1	f_2	f_3
m_1	5.166	5.579	4.603	5.270	5.668	4.693
m_2	5.358	5.823	4.828	5.376	5.885	4.874
m_3	4.672	4.972	4.186	4.800	5.078	4.290

while for the males the payoffs are:

Male Payoffs

	NE			CPO		
Male	f_1	f_2	f_3	f_1	f_2	f_3
m_1	4.992	5.195	4.586	4.994	5.195	4.586
m_2	5.371	5.618	4.840	5.434	5.618	4.856
m_3	4.584	4.748	4.333	4.585	4.749	4.334

In this example, in general the payoff for man i associated with marrying woman j varies little whether the household behaves efficiently or not. This is due to the fact that, given the state variables and notional bargaining power parameter value $\alpha = 0.5$, in most potential marriages the participation constraint is binding for the male. In such a situation, the value of the marriage to the male is the same under NE and CPO. As a result, the women are receiving most or all of the surplus generated by efficient household behavior, and consequently their payoffs from any marriage are larger under CPO than under NE.

In the example, and in the empirical work which follows, we will assume that the common distribution of match draws, F, is a mean 0 normal with standard deviation σ_ε. In the case where $\sigma_\varepsilon = 0$, there are no match-specific shocks, since the distribution of F is degenerate with all of its mass concentrated at 0. Any strictly positive value of σ_ε produces a nondegenerate continuous distribution of the match-specific preference shock. In performing the computations reported below, we simply report the proportion of draws for which the given marriage pattern was observed out of 100,000 draws of the $\{\varepsilon_{ij}\}_{i=1,j=1}^{3,3}$ required for the computation of the payoff matrices under NE and CPO behavior.

There are six possible marriage patterns that are possible with three males and three females, assuming that everyone is matched, and they are listed in the first column of Table 1. The second and third column of the table correspond to the case for which there is no random component of match quality. We see that in this case both NE.and CPO imply the same marriage pattern. In columns 3 and 4, we see that by adding random match quality the situation changes markedly. In particular, the marriage pattern that existed with probability 1 with no match heterogeneity now occurs with probability 0.927 under NE and with probability 0.862 under CPO. The second most probable pattern occurs with likelihood 0.071 under NE and with likelihood 0.136 under CPO. Thus, even when switching modes of behavior does not change the marriage market prediction when there is no match heterogeneity, it does in the presence of a random match component. This is due to the fact that switching modes of behavior changes the sizes of the differences in the "deterministic" component of payoffs, which results in changes in marriage patterns under the two R.

The last two columns of the table report the marriage matching results when after we further increase the spread in the distribution of the random component. Such increases give the random match component a more important role in determining the final payoffs. Since that the random match values are i.i.d. draws, this will lead to more dispersion in the marriage pattern distribution. When we double the size of the standard deviation σ_ε, we see that the proportion of times that the modal marriage pattern is observed declines to 0.640 under NE and 0.587 under CPO. The second most frequently observed marriage pattern has increased in likelihood to 0.275 under NE and 0.328 under CPO. Now every type of marital pattern has a probability of being observed of at least 0.001.

This example has served to illustrate how differences in marital sorting patterns can potentially be utilized to understand how spouses coordinate their decisions within existing marriages. We now turn to a consideration of how more formal statistical procedures can be constructed to examine this question.

3 Econometrics

We consider the estimation of the marriage market equilibrium in a sequential fashion. We begin with the issue of the estimation of (G_1, G_2), the distribution of gender types, with G_1 being the population cumulative distribution function of male characteristics and G_2 the analogous c.d.f of female characteristics. In this paper we do not treat the difficult censoring issues that arise when not all household members supply time to the labor market or in household production. In such a situation, nonparametric point identification of household members' characteristics will be impossible and the particular identification and estimation strategy we employ will not be available. Given that there are no corner solutions in the time allocation decisions within the household, we are able to posit that the entire vector

$$A_k = (h_{1k}, h_{2k}, \tau_{1k}, \tau_{2k}, w_{1k}, w_{2k}, y_{1k}, y_{2k}),\ k = 1, ..., N,$$

is observable by the analyst. For the present, we have constructed the male and female indexing so that in the data male i is married to female i, $i = 1, ..., N$, though in the following subsection on marital sorting we will not impose this indexing convention. It will be useful to partition this vector into two subvectors,

$$\begin{aligned} A_k^1 &= (h_{1k}, h_{2k}, \tau_{1k}, \tau_{2k}), \\ A_k^2 &= (w_{1k}, w_{2k}, y_{1k}, y_{2k}), \end{aligned}$$

with A_k^1 representing the (endogenous) time allocations of household k and A_k^2 the observable (to the analyst) state variables. The unobservable state variables in household k are $(\lambda_{1k}, \lambda_{2k}, \delta_{1k}, \delta_{2k})$. Then denote the remaining unobserved household characteristics by

$$A_k^3 = (\lambda_{1k}, \lambda_{2k}, \delta_{1k}, \delta_{2k}).$$

The data used in the empirical work discussed below are drawn from the Panel Study of Income Dynamics (PSID). In keeping with the static setting of the model, we use data pertaining to household characteristics and time allocation decisions in one year, 2006. We chose this year because information on the time spent in household tasks is widely available for both spouses in that year and is relatively recent.

We assume that the PSID is randomly drawn from the population distribution of married households in this year (which is an unlikely situation, admittedly). Since we have a random sample of households, we also have a random sample of household members *given* the marriage assignment rule.

Using a random sample of N households from the population, the first task is to estimate the distribution functions G_1 and G_2. For household k, we can restate (1) as

$$A_k^1 = R(A_k^2 \cup A_k^3).$$

Proposition 7 *Assume all households in the population behave according to R, and that R is invertible in the sense that there is a unique value of A_k^3 such that*

$$A_k^3 = R^{-1}(A_k^1 \cup A_k^2) \tag{6}$$

for all values of $A_k^1 \cup A_k^2$. Then the distributions G_1 and G_2 are nonparametrically identified and can be consistently estimated.

Proof: Given knowledge and invertibility of R, then R^{-1} is a known function. If A_k^1 and A_k^2 are observed without error, then the vector A_K^3 is observable as well. Since the vectors A_k^1 and A_k^2 are observed for a random sample of households, then A_k^3 is as well. Define the vectors

$$\begin{aligned} X_k &= (A_k^3, w_{1k}, w_{2k}, y_{1k}, y_{2k}), \\ X_k^1 &= (\lambda_{1k}, \delta_{1k}, w_{1k}, y_{1k}), \\ X_k^2 &= (\lambda_{2k}, \delta_{2k}, w_{2k}, y_{2k}). \end{aligned}$$

The vector X_k^1 is an i.i.d. draw from G_1 and X_k^2 is an i.i.d. draw from G_2. Then define

$$\begin{aligned}\hat{G}_1^N(x) &= N^{-1}\sum_{k=1}^{N}\chi(X_k^1 \le x), \\ \hat{G}_2^N(x) &= N^{-1}\sum_{k=1}^{N}\chi(X_k^2 \le x).\end{aligned}$$

Since $\{X_1^1, ...X_N^1\}$ and $\{X_1^2, ..., X_N^2\}$ are both random samples from their respective populations, we know that

$$\operatorname*{plim}_{N\to\infty} \hat{G}_i^N(x) = G_i(x),\ i = 1, 2,$$

by the Glivenko-Cantelli Theorem.■

We emphasize the important point that the mapping from the observed subset of state variables and household decisions to the unobserved state variables is independent of any random match values. This follows from Proposition 6, which established that the actions taken under rule R were only a function of the state variables A_{1k} and A_{2k}, and not the match-specific heterogeneity draw. The inverse function inherits this property.

The following important implication follows.

Proposition 8 *Let $\Re$ be the set of rules that determine time allocations in the household and that are invertible in the sense of (6). Then all $R \in \Re$ are equivalent descriptions of sample information.*

Proof: Consider a household k in the sample. We observe four household choices $A_k^1 = (h_{1k}, h_{2k},\tau_{1k},\tau_{2k})$ and we have four unobservable characteristics of the spouses. Thus given any $A_k^2 = (w_{1k},w_{2k},y_{1k}, y_{2k})$ and any $R \in \Re$, there exists a unique vector of characteristics $(\lambda_{1k}, \lambda_{2k}, \delta_{1k}, \delta_{2k})$ that generate A_k^1, or

$$A_k^1 = \Gamma(\lambda_{1k}(R), \lambda_{2k}(R), \delta_{1k}(R), \delta_{2k}(R)|A_k^2, R).$$

Then for any two $R, R' \in \Re$, $R \neq R'$,

$$\begin{aligned}&\Gamma(\lambda_{1k}(R), \lambda_{2k}(R), \delta_{1k}(R), \delta_{2k}(R)|A_k^2, R) \\ =\ &\Gamma(\lambda_{1k}(R'), \lambda_{2k}(R'), \delta_{1k}(R'), \delta_{2k}(R')|A_k^2, R'),\end{aligned}$$

which describes a correspondence between $(\lambda_{1k}, \lambda_{2k}, \delta_{1k}, \delta_{2k})(R)$ and $(\lambda_{1k}, \lambda_{2k}, \delta_{1k}, \delta_{2k})(R')$.

Consider any distance function

$$\mathbb{Q}(A_k^1, \hat{A}_k^1(\lambda_{1k}, \lambda_{2k}, \delta_{1k}, \delta_{2k}|A_k^2, R)),$$

where $\hat{A}_k^1$ is the predicted value of the household time allocations given the characteristics $(\lambda_{1k}, \lambda_{2k}, \delta_{1k}, \delta_{2k})$, A_k^2, and R. Given invertibility

$$\begin{aligned}&(\lambda_{1k}(R), \lambda_{2k}(R), \delta_{1k}(R), \delta_{2k}(R)|A_k^2, R) \\ =\ &\arg\min_{\lambda_{1k},\lambda_{2k},\delta_{1k},\delta_{2k}} \mathbb{Q}(A_k^1, \hat{A}_k^1(\lambda_{1k}, \lambda_{2k}, \delta_{1k}, \delta_{2k}|A_k^2, R))\end{aligned}$$

and

$$\mathbb{Q}(A_k^1, \hat{A}_k^1((\lambda_{1k}(R), \lambda_{2k}(R), \delta_{1k}(R), \delta_{2k}(R)|A_k^2, R)|A_k^2, R)) = 0,$$
$$\forall\ R \in \Re$$

■

Because of the flexible parameterization of spouses in terms of their types, if $\Re$ contains more than one element there are multiple ways to "reparameterize" the data, in essence. There are four pieces of data for each household (where data are considered to consist of observations on endogenous variables only) and there are four unknown parameters for each household. Thus model fit must be perfect. In the language of statistics, such a model is termed *saturated,* and it is well-known that there are typically a number of equivalent ways in which such a model can be parameterized, with all parameterizations being equivalent in the sense of being perfectly consistent with the data. We have made this point by defining a set of behavioral rules, $\Re$, that the household can follow but that produce identical (observed) choices given the subset of state variables that are observable.

The cardinality of $\Re$ depends on assumptions made regarding the functional form of the utility and household production functions and the features of the data. In a companion paper (Del Boca and Flinn, 2012) we prove the following.

Proposition 9 *For the Cobb-Douglas specification of individual preferences and household production technology and assuming no corner solutions, the Nash Equilibrium (NE) and Constrained Pareto Optimal (CPO) models of household decision-making are elements of* $\Re$.

This proposition carries the important implication that it is not possible to determine whether household decisions are made according to Nash Equilibrium or in the Constrained Pareto Optimal setting by observing only within-household behavior. This "impossibility" result is due to the flexible specification of population heterogeneity utilized. By restricting the variability of these underlying parameters in the population, it will be possible to develop tests pitting the two allocation rules against one another. However, the outcome of such a test will be heavily dependent upon the parametric restrictions adopted.

We note that the allocation mechanism consistent with efficient household behavior, CPO, is actually underidentified in the sense that the notional, or ex ante, Pareto weight α is not identified. We consider this behavioral model as belonging to $\Re$ for a predetermined value of α. The reader should bear in mind that there are actually a continuum of allocation mechanisms associated with CPO for all of the notional values of $\alpha \in (0, 1)$. In the empirical analysis conducted below we focus on the "symmetric" case of $\alpha = 0.5$.

3.1 Marital Sorting

Flexible specifications of population heterogeneity reduce the analyst's ability to derive distinguishable empirical implications from elements of a class of modes of behavior. However, they do provide possibilities for developing tests (of some form) based on marital sorting patterns.[8] We explore the construction of such a test in this subsection.

The main problem to be faced in examining marriage sorting patterns is the definition of the spousal choice set for each individual in the population being studied. When individuals make marriage decisions, their choices often may be limited to a small set of individuals (pop stars and professional athletes being notable exceptions). The analyst will typically have no information regarding the individual marriage markets of husbands and wives interviewed in national surveys such as the PSID, making it difficult to assess the likelihood that this pair of individuals would have chosen one another under the household allocation rule R. We have decided to make the nonobservability of the marriage market and its random nature a key component of the measure of goodness of fit we propose below.

We recognize that the random marriage market assumption, which is somewhat akin to the assumption of a random arrival time of marriage candidates in a search-based model of marriage, sidesteps the important issue of endogenous marriage market formation. For example, by attending a college away from their parental home, young adults potentially enter into a much larger marriage market with a different composition of possible partners than would be the case if they were to not attend college and not move away from their hometown. In the empirical analysis discussed below, we will ignore these endogeneity issues, though they are potentially extremely important for explaining marital sorting patterns.

Given the nature of the marriage equilibrium concept we are using, and ignoring the issue of uniqueness (we only consider the male-proposer stable matching), we have assumed that the households in our N household sample from the PSID are a representative sample from the population of all married couples. We order the married males and females within the sample in an arbitrary manner m_1 through m_N denoting the males and $f_1, ..., f_N$ for the females. The function Γ_D associates the index of a wife in the data with the index of her husband as $f_{\Gamma_D(i)}$, where i is the husband's index. Thus $66 = \Gamma_D(1)$ indicates that female 66 in the sample is married to male number 1. Without loss of generality, we index the spousal pairing s_i to be consistent with the index of husband. Thus, in the data, $s_i = \{m_i \; f_{\Gamma_D(i)})\}$, $i = 1, ..., N$.

In our view it is problematic to think of all individuals in the PSID subsample with which we work as belonging to the same marriage market. Instead, we think of marriage markets as being comprised of individuals selected from a population of potential marriage partners. For purposes of the empirical analysis we have to impose some strong restrictions on the manner in which these marriage markets are formed. A marriage market will be denoted by σ, and includes an equal number of males and females. A marriage market of size n is given by

$$\sigma = \{m_{(1_m)}, m_{(2_m)}, ..., m_{(n_m)}; f_{(1_f)}, f_{(2_f)}, ..., f_{(n_f)}\},$$

[8]Marital sorting is but one phenomenon that could be used to distinguish between modes of intrahousehold behavior. Others include divorce decisions and investments in marriage-specific capital.

with (j_g) indicating the index of the male or female with the j^{th} smallest index among their gender g in the marriage market. We will denote the marital partner of male (i_m) in marriage market σ by $(j_f) = \Gamma_\sigma((i_m))$, $i_m = 1, ..., n$.

We will assume that any marriage market σ is comprised of a closed set of married individuals, where "closed" is used in the following sense.

Definition 10 *A marriage market σ is closed if all $m \in \sigma$ have the same potential marital partners $f \in \sigma$ and no others and if all $f \in \sigma$ have the same potential marital partners $m \in \sigma$ and no others.*

Example 11 *A marriage market with three members of each sex is given by*

$$\sigma = \{m_{17}, m_{25}, m_{109}; f_7, f_{76}, f_{99}\}.$$

Each male has the females 7, 76, and 99 in their personal marriage market and no others. Each female has only the males 17, 25, and 109 in their personal marriage market. If under rule R the male-proposer stable matching has male 17 married to female 99, male 25 married to female 7, and male 109 married to female 76, then $\Gamma_\sigma(17) = 99$, $\Gamma_\sigma(25) = 7$, *and* $\Gamma_\sigma(109) = 76$. *The spousal pairs are* $s_{17} = \{m_{17}\ f_{99}\}$, $s_{25} = \{m_{25}\ f_7\}$, *and* $s_{109} = \{m_{109}\ f_{76}\}$.

Thus in the example above, $(1_m) = 17$, $(2_m) = 25$, $(3_m) = 109$ and $(1_f) = 7$, $(2_f) = 76$, and $(3_f) = 99$. Even within our very stylized conceptualization of the marriage market, there is no reason to expect that all closed marriage markets are of the same size. We will repeat our empirical analysis assuming different-sized marriage markets, ranging from $n = 2, ..., 7$, but we will be assuming that in each case *all* marriage markets are of the same size. Since we will are not able to observe any actual marriage market, we thought it would be gratuitous to allow for a world consisting of a mixture of marriage market sizes, though clearly this is a potentially interesting avenue to pursue.

Definition 12 *A size n closed marriage market σ_n is constructed by taking n draws without replacement from the populations $\tilde{G}_1$ and $\tilde{G}_2$, where $\tilde{G}_j$ is the augmented type distribution associated with gender j.*

The "augmented" type distribution is important to the bootstrapping procedure we implement. This distribution is defined with respect to not only the state variables that determine household behavior in the event of a marriage, but also the draws of match values associated with each potential marriage partner in the global marriage market. This allows for the presence of identical males and/or females in terms of the state variables that determine household actions in the same marriage market (i.e., S_{ij}), while requiring that all potential matches differ in their idiosyncratic match values. For example, say that male m_{13} and m_{21} have the same characteristics in terms of λ, δ, w, and Y. They are considered distinct marriage partners as long as their idiosyncratic match draws with all potential females in the marriage market are different. While the "economic" payoff is the same for any female in the marriage market (under either rule R), the total payoff is not as long as the psychic payoff associated with being in a marriage with either is different for each female in the market. Since the psychic payoff is continuously distributed, this is the case with probability 1.

This assumption regarding the construction of labor markets implies that marriage markets σ of any size n are defined in a purely exogenous manner with respect to the behavioral rule R assumed to govern within-household interactions. This strict exogeneity assumption is important for the construction of our comparisons of the fit among the behavioral rules we evaluate.

It is important to emphasize that we do not view "local" marriage markets as being (random) partitions of some fixed set of individuals in search of mates at given point in time. Instead, marriage markets are constructed from random draws from distributions $\tilde{G}_1$ and $\tilde{G}_2$ up to the size of the market, n. When there are a continuum of males and females in the population of potential marriage market members and when the type c.d.f.s are absolutely continuous, then the assumption that sampling be without replacement is unnecessary since the probability of the same individual type being drawn more than once in any finite marriage market σ_n is zero. For the case in which the male and female type distributions are discrete and there is one realization of the match-specific draws,[9] then there is a strictly positive probability that a given type male will simultaneously be a member of several marriage markets of size n. Our assumption that in any given marriage market σ_n no more than one of any type of male or female is present is substantive. The purpose of the assumption is to rule out ties when implementing the GS algorithm. This issue will come up again when we discuss bootstrapping our metric of sample fit.

3.2 Choosing Between R

Our objective is to compare the relative consistency of the marriage patterns implied under NE and CPO with those observed in the PSID subsample with which we work. We begin by discussing some of the problems one must face in constructing a reasonable metric and the case against some which are implementable but very much ad hoc. We will then derive a measure that we will employ in the empirical analysis and which has a more solid foundation and interpretation from a statistical point of view.

The data set with which we work contains 282 married individuals. If we assume that there is no measurement error in the data, then under rule R we can "back out" all of the

[9] The discrete case is the relevant one empirically, since our estimators of the type distributions G are empirical distribution functions with a mass of N^{-1} at each of the values of the gender-specific state variables directly observed or inferred under rule R. Even for the augmented distributions, those that include match-specific heterogeneity, the draws are assigned at the population level (the population from which the size n marriage markets are constructed), so that there exists a finite number of types in this case as well.

state variables for the males and females in the sample. Under the assumption that they comprised one unified marriage market (i.e., a closed marriage market of size $n = 282$), we can apply the GS algorithm and compare the marriage patterns under $R = NE$ and $R = CPO$. It will come as no surprise that the spousal pairings predicted under either rule will be inconsistent with those observed in the set of 282 matches. Since everything is deterministic (given the rule R), there is no scope for comparing the NE and CPO rules - both are simply incorrect - meaning, the actual sorts are probability zero under the model.

One can begin to attempt to add randomness to the stable matchings, as was done in a previous draft of this paper, by assuming measurement error in some of the state variables. We had assumed that wages were measured with error and used estimates of the measurement error process for wages in the PSID taken from the validation study of Bound et al. (1994). Even with measurement error in wages, the likelihood of drawing a "true" wage vector for males and females that would imply the 282 observed marriages under either rule is for all practical purposes equal to zero. Thus if the goal is to utilize such a sharp prediction metric, which records a success only when all marriages are correctly predicted, it is essential to limit the size of marriage markets. This is one element of the metric designed below.

Of course, there are other "softer" prediction metrics that could be used. One that was employed by us earlier was one version of an assortative mating metric which included only the correlation of the wages of husbands and wives (recall that a wage measure was available for everyone due to our selection criterion that all sample members be working). By assuming measurement error in wages, given a draw of the measurement error vector, one can compute the stable matching and the correlation in the wages of spouses under a behavioral rule R. This correlation can be compared with the correlation found in the data. Because the true wage is assumed to be a random variable, the wage correlation implied under rule R is also a random variable, and bootstrap methods can be used to assess whether or not the observed correlation of wages in the sample is unsurprising from the point of view of the distribution of the implied correlation under a given R. This metric has the advantage of being easily computable under any size marriage market, though, as is true for any model of marital sorting, its validity will crucially depend on the correctly specifying the participants in the marriage market. Moreover, a behavioral rule may fare poorly using a hard prediction metric (such as the proportion of matches successfully predicted) but may do well in terms of the wage correlation metric.[10]

We develop a metric of fit that can also be used as a basis for the estimation of parame-

[10] The case for using a wage correlation metric, for example, is strongest when considering marriage patterns in a transferable utility framework. Say that male and female types (e.g., wages) are givern by a and b, and that the productivity of the marriage is a superadditive function of spousal types. In this case the female matched to male a, $b^*(a)$, will be a strictly increasing function of a, which means that there is perfect (positive) assortative mating. In our case, the value of the marriage to each individual is a function of a number of state variables, not solely wages, so that examining association on one dimension, such as wages, yields only a partial reading on the sorting process.

ters or other objects that are not identified purely from within household time allocation data; we will discuss some of these possibilities below. We consider a spousal pair i from the PSID, where $i = 1, ..., N$. For a given marriage market of size n, what is the probability that we would have found i married to female $\Gamma(i)$? Clearly it is zero unless both male $i \in \sigma$ and female $\Gamma_D(i) \in \sigma$, so for the probability to be positive both of these individuals must belong to this closed market of size n. Let $\sigma_{-i}(n)$ denote the set of individuals in the marriage market of size n other than male i and female $\Gamma_D(i)$. These individuals are drawn randomly (without replacement) from the augmented type distribution of males and females, $\tilde{G}_1$ and $\tilde{G}_2$ respectively (the without replacement restriction also applies to the included spouses i and $\Gamma_D(i)$ in that another male of type i cannot be included in σ and neither can another female of type $\Gamma_D(i)$). With the randomly constituted marriage market (conditional on the inclusion of male i and female $\Gamma_D(i)$), the GS algorithm is applied given the preference orderings generated by (M_σ, F_σ, R), where M_σ denotes the characteristics of the males included in market σ and F_σ the characteristics of the females in that market. These characteristics include the random match-specific heterogeneity draws that are relevant for all the potential pairings in the marriage market. We are interested in the event in which $\Gamma_\sigma(i) = \Gamma_D(i)$, i.e., the event in which the spouse observed in the data is the same as the spouse observed in marriage market σ.[11]

Now in any given marriage market of size n that includes i and $\Gamma_D(i)$, we will either have $\Gamma_\sigma(i) = \Gamma_D(i)$ or not. Then we will define a conditional probability function

$$P(\Gamma_\sigma(i) = \Gamma_D(i)|(m_i + \varepsilon_{i,\Gamma_D(i)},\ f_{\Gamma_D(i)+}\varepsilon_{i,\Gamma_D(i)}, \sigma_{-i}(n,\varepsilon)), R) = \begin{cases} 1 & \text{if} \quad \Gamma_\sigma(i) = \Gamma_D(i) \\ 0 & \text{if} \quad \Gamma_\sigma(i) \neq \Gamma_D(i) \end{cases}$$

which is in fact just an indicator function assuming the value 0 or 1 since we have conditioned on marriage market composition. Let $\tilde{G}_{-i}(\sigma_{-i}(n), \varepsilon)$ denote the probability distribution over marriage market members (males and females) other than male i and female $\Gamma_D(i)$ for a given realization of match-specific shocks ε. Then the marginal probability of observing male i married to female $\Gamma_D(i)$ is

$$\begin{aligned} P(s_i^D|R, n) \quad &= \quad \int\int P(\Gamma_\sigma(i) = \Gamma_D(i)|(m_i + \varepsilon_{i,\Gamma_D(i)},\ f_{\Gamma_D(i)} + \varepsilon_{i,\Gamma_D(i)}, \hat{\sigma}_{-i}(n,\varepsilon)), R) \\ &\quad \times d\tilde{G}_{-i}(\sigma_{-i}(n,\varepsilon))dF_N(\varepsilon) \qquad (7) \end{aligned}$$

where $F_N(\varepsilon)$ is an N-variate distribution (which under our independence assumption is the product of N univariate distributions F). In general, for a fixed n and a given R,

[11] This procedure is similar in spirit to and partially inspired by the choice-set sampling procedure developed by McFadden (1978) for the estimation of a model of residential location choices. In his case, their was no two-sided matching feature of the model (locations could accommodate any resident who desired to live there) and the logit structure implied an IIA property. In that case, the choice made would be the same no matter which subset of alternatives was available. The same is not true regarding stable matches. We thank Al Roth for clarifying our thinking on this point.

$P(s_i^D|R,n) \in (0,1)$. This probability is simply the measure of potential marriage markets of size n that contain male i and female $\Gamma_D(i)$ and that would have produced a marriage between them under rule R.

3.3 Sample Likelihood Construction

Consider the sequence of terms $P(s_i^D|R,n)$, $i = 1,...,N$. These are all conditional probabilities, and under our interpretation of the way in which marriage markets are formed, $P(s_i^D|R,n)$ is independent of any other spousal pairing s_j^D, $j = 1,...,n;\ j \neq i$. Thus, the probability of the set of spousal pairings $(s_1^D,...,s_N^D)$ is the product of probabilities of each of the pairings, or

$$P(s_1^D,...s_N^D|R,n) = \prod_{i=1}^{N} P(s_i^D|R,n).$$

This of course is just the likelihood of observing the pairings s_i, $i = 1,...,N$, or

$$L(R,n) = \prod_{i=1}^{N} P(s_i^D|R,n),$$

with the log likelihood given by

$$\ln L(R,n) = \sum_{i=1}^{N} \ln P(s_i^D|R,n).$$

The most noteworthy feature of this as a likelihood function is that no parameters appear explicitly. Instead, the likelihood is solely a function of the distribution of male and female characteristics, $\tilde{G}_1$ and $\tilde{G}_2$, which have been assumed known to this point (conditional on a realization of ε), and the behavioral rule R.

In constructing the partial log likelihood function, we have assumed that the probability of any given spousal matching is independent of the others in the sample. This followed from the assumption that $\tilde{G}_1$ and $\tilde{G}_2$, used to construct the distribution $\tilde{G}(\sigma_{-i}(n,\varepsilon))$ for each i, were known in addition to the characteristics of the given spousal pair i. In terms of the empirical application, they are not in fact, and are instead estimated based on the state variables observed and inferred from the within-household time allocations *under the assumption of a given rule* R. In addition, in order to form an estimate of $\tilde{G}$ we need to know the distribution of match-specific heterogeneity, F. As in the extended example in Section 2.4, we assume that F is a mean 0 normal distribution with standard deviation σ_ε. While in theory it is possible to estimate σ_ε, in terms of computational time it is not practical to do so.[12] As a result, we have simply assumed that $\sigma_\varepsilon = 0.1$.[13] Then we

[12] For a given value of σ_ε, it takes approximately one week to compute the log likelihood ratios for the four regions and the total sample for marriage market sizes $n = 2,...,7$ using three workstations and exploiting time savings from parallel processing. Thus iterating over values of σ_ε is not a practical alternative.

[13] We calculated some log likelihood ratios using a few alternative values of σ_ε and found that the inferences

use $\hat{G}_1(R)$ and $\hat{G}_2(R)$, which are the first-stage estimates of the distributions of the state variables ($\lambda\ \delta\ w\ Y$) for men and women using the mapping associated with rule R, to determine the estimated conditional (on ε) distribution $\hat{G}_{-i}(\tilde{\sigma}_{-i}(n,\varepsilon)|R)$, so that

$$\hat{P}(s_i^D|R,n) = \int P(\Gamma_\sigma(i) = \Gamma_D(i)|(\hat{m}_i(R) + \varepsilon_{i\Gamma_D(i)}, \hat{f}_{\Gamma_D(i)}(R) + \varepsilon_{i\Gamma_D(i)}, \tilde{\sigma}_{-i}(n,\varepsilon)), R) \times d\hat{G}_{-i}(\tilde{\sigma}_{-i}(n,\varepsilon)|R)d\Phi_N(\varepsilon;\sigma_\varepsilon^2),$$

where Φ_N is an N-variate multinormal distribution with mean vector 0 and covariance matrix $\sigma_\varepsilon^2 I_N$ and where $\hat{m}_i(R)$ and $\hat{f}_{\Gamma_D(i)}(R)$ signify that a subset of the characteristics of the spouses in the i^{th} couple are estimated using the mapping associated with behavioral rule R. Then the estimated sample log likelihood is given by

$$\ln \hat{L}(R,n) = \sum_{i=1}^{N} \ln \hat{P}(s_i^D|R,n).$$

There is what we would characterize as "weak" dependence between probabilities of different spousal matchings that is transferred through the estimated distributions of state variables. That is, the estimated conditional (on ε) probability distribution of marriage markets in which couple $(i, \Gamma_{D(i)})$ is potentially embedded is determined by the sample composition of households. As the size of the potential marriage market population, N, increases, the dependence in these probabilities becomes weaker as $\text{plim}_{N\to\infty} \tilde{G}_{-i}(\tilde{\sigma}_{-i}(n,\varepsilon)|R) = G_{-i}(\sigma_{-i}(n,\varepsilon))$ at all points of continuity of G for a given rule R and for all draws ε. In the construction of this log likelihood function we ignore this dependence.

We note that the log likelihood function $\ln \hat{L}(R,n)$ depends on no unknown parameters in the conventional sense of the term "parameter," at least when we consider the distribution F to be known. However, one can think of posing the problem of determining the rule most consistent with the GS male-proposer stable matching within the class of all feasible allocation rules, which we denote by Ω.[14] Then the maximum likelihood estimator of

$$\hat{R}(n) = \arg\sup_{R\in\Omega} \ln \hat{L}(R,n).$$

The manner in which one could define the set Ω in order to make this estimator operational is not immediately obvious, but is a potentially interesting exercise.

drawn were not affected.

[14]When we know the distribution of individual types in the population of potential marriage market members it is not necessary to estimate the type distribution in the first stage, and when we consider the set of all possible decision rules it is not necessary to limit consideration to those that deliver consistent first stage estimates of the type distributions. Note that $\Re \subset \Omega$, since the class $\Re$ are feasible allocation rules that are also invertible in the sense of (6) and we know that there exist feasible allocation rules that are not invertible in this sense (an example of one can be found in Del Boca and Flinn (2012)). Feasibility here and in the text simply means that the allocation rule must deliver outcomes that lie within the household's choice set.

On a more practical level, $\ln \hat{L}(R, n)$ could in principle be used to estimate more conventional parameters, such as n or the notational Pareto weight α from the CPO behavioral rule for a given rule R. A moment's reflection reveals that n cannot be estimated in any meaningful sense since setting $n = 1$ (implying that only one's spouse was in the choice set) yields the maximum value of the log likelihood value of 0 independently of the rule R or of the characteristics of sample members.[15]

Estimation of the notational Pareto weight is a different matter, however. We have performed the empirical analysis involving the CPO rule under the assumption that $\alpha = 0.5$. Using only data on intrahousehold time allocation we showed that α was not identified when G_1 and G_2 were estimated (and identified) nonparametrically. However, using the marital sorting information this is no longer the case. In particular, we can define the log likelihood function $\ln \hat{L}(CPO, \alpha, \sigma_\varepsilon, n)$ and define a maximum likelihood estimator of α as $\hat{\alpha}(\sigma_\varepsilon, n) = \arg\sup_{\alpha \in (0,1)} \ln \tilde{L}(CPO, \alpha, \sigma_\varepsilon, n)$. This is an interesting approach to identifying the Pareto weight, but we have not pursued it here given our interest in comparing these rules on more equal terms.[16]

Our focus is on evaluating $\ln \hat{L}(R, n)$ at two different rules, NE and CPO (with $\alpha = 0.5$), for a sequence of values of n. As this discussion hopefully has made clear, comparing NE and CPO is not a hypothesis testing exercise, it is merely a question of fit. We think of $\Re$ as representing a relatively abstractly-defined set of allocation rules, with NE and CPO representing two members of that set. The rule that is superior for a given marriage market size n is simply the point associated with the largest value of $\ln \tilde{L}(R, n)$. The difference in log likelihood values between the two rules is

$$\ln \hat{L}(CPO, n) - \ln \hat{L}(NE, n) =$$
$$\sum_{i=1}^{N} \ln \hat{P}(s_i^D | CPO, n) - \sum_{i=1}^{N} \ln \hat{P}(s_i^D | NE, n)$$

Given a value of n and the sample, CPO is to be preferred to NE when the value of this expression is positive, for when this is the case the sample matches are more likely under CPO.

[15] In other words, the partial log likelihood function is monotone in the parameter n implying maximization on the boundary of the parameter space.

[16] There is potentially another problem with the estimation of α in the manner proposed here. As $\alpha \to 1$ from below or $\alpha \to 0$ from above, all households are constrained to adjust the notional, or ex ante, weight α to be able to guarantee a welfare level at least as high as that attained under NE to both household members. In such a case the notational weight α can not be point-identified and instead can only be set-identified.

A more practical problem, as was pointed out when we discussed the possibility of estimating σ_ε, is the computational time required in maximizing such a complex log likelihood function.

3.4 Bootstrapping

In order to gain an idea of the sampling distribution of our comparative measure of fit, the log likelihood ratio, we utilize bootstrapping methods. The presence of the random match-specific shocks is critical for our being able to employ these methods. Consider the entire marriage market from which our actual size n marriage markets are drawn. Since our sample consists of married couples, the bootstrap estimate of the sampling distribution of the log likelihood ratio is based on drawing husband and wife pairs from the PSID. In any bootstrap sample that is different from the actual PSID sample, some pairs will be represented more than once. This means that in some size n marriage markets there is a nonzero probability that the same male or the same female will appear, which will result in ties in the preference orderings leading to indeterminacies in the set of stable matches resulting from the application of the Gale-Shapley algorithm. Tie-breaking conventions are inherently arbitrary, and using them would introduce a degree of indeterminacy into the log likelihood values we compute.

In a bootstrap sample of size N from the original sample of size N, the presence of repeated observations of couples in the bootstrap sample is essentially assured. Random match-specific heterogeneity serves to make the payoffs associated with marrying a potential partner distinct even if there exist other potential partners who possess the same values of the state variables $(\lambda\ \delta\ w\ I)$. For example, say that in a marriage market male i can potential match with female j or female j' who are observationally equivalent. Under any given rule R, male i will prefer female j if and only if $\varepsilon_{ij} > \varepsilon_{ij'}$. Since the ε are i.i.d. draws from an absolutely continuous distribution, the probability of indifference is zero.

There is a large computational burden associated with this empirical work. After selecting a market of size n from the larger marriage market, all "deterministic" payoffs associated with each potential pairing are computed, with the random draws ε then included to determine the total gender-specific payoffs associated with the each pairing then utilized to determine the GS outcomes for the marriage market. This process is repeated for each individual in the sample, across draws of the ε, across the two rules R, and then across the number of bootstrap replications, B. This is a time-consuming process, even though we are not estimating any parameters.

We give a brief description of the process to provide some clarity. The number of bootstrap replications, B, was set at 20. This is sufficient to give us some idea of the distribution of the log likelihood across random samples in the population. For a given bootstrap sample, we then drew a vector of random match-specific shocks ε from the i.i.d. mean 0 normal with standard deviation $\sigma_\varepsilon = 0.1$ for all potential matches between males and females in the bootstrap sample. For each bootstrap sample, we drew 10 values of the vector ε. From these, we were able to compute the payoff matrices for all males and females in the bootstrap sample of size N under behavioral rule R. Then for each sample married couple, we drew 20 marriage markets of size n, and computed the proportion of times the sample couple were married to each other in the 200 (20 marriage markets of size $n \times 10$

different ε draws) potential marriage markets in which they found themselves.[17,18] The log likelihood in bootstrap sample b and under rule R was then the sum of these log likelihoods over the N couples in the sample. We have 20 values of the log likelihood computed under CPO and NE, and examine the distributions of the log likelihood ratios.

3.5 Sample Information and Test Power

The test we have defined above only works if the marriage predictions under the two competing (simple) hypotheses are actually different. In this subsection, we provide some evidence that the sample information is sufficient to conduct a meaningful test to determine which of the behavioral models is more consistent with the observed marital sorts.

Our discussion proceeds via an example that is closely related to a small set of the results reported below. In this example, we assume the absence of random match-specific heterogeneity, so that $\sigma_\varepsilon = 0$. We only consider marriage markets of size two. In terms of evaluating alternative matches, for each of the (married) couples in our sample (described below), we utilize only other married couples. Since there are 282 married couples in the sample, we determine the proportion of times each couple would have been married to each other if in a two-male and two-female marriage market with each of the other 281 couples serving as the alternative male and female. We compute the proportion of times they would have been correctly matched under CPO and the proportion of times they would have been correctly matched under NE. If these proportions are identical, the couple provides no information on the household behavioral rule.[19] If the difference in these proportions is "large," then this couple provides a significant amount of information regarding the underlying behavioral rule.

Figure 1 contains the histogram of the difference between the proportion of correct marriage predictions under CPO and under NE for the 282 couples in the sample. The first thing to note is that there is considerable dispersion in this distribution. For only 10 couples out of the 282 are the proportion of correct predictions under CPO and NE

[17] The marriage markets of size n for couple i are constructed by sampling without replacement from the other $N-1$ males and $N-1$ females in the marriage market, not including themselves. An alternative male $j \neq i$ could not be included more than once in the alternative $n-1$ males for female i since this would lead to ties in the GS algorithm. Note that by not being the same, we mean that they can not be the same both in terms of the $(\lambda\ \delta\ w\ Y)$ and the ε draw; If two members of the population have the same values of the vector $(\lambda\ \delta\ w\ Y)$ then it must be the case that their respective ε draws are different..

[18] For the case in which the couple were never married to one another under rule R in any of the 200 marriage markets we constructed, we assigned a correct match proportion of 0.001, since the proportion must be greater than 0 for the log likelihood to be defined. Clearly, by increasing the number of ε draws and the number of marriage market draws we would be able insure that this proportion is strictly greater than 0 for all couples in the sample for either rule R, but computational costs restricted our ability to do so.

[19] Note that by considering cases in which $\sigma_\varepsilon > 0$, every case will be informative for distinguishing between the two behavioral rules. Our choice of $\sigma_\varepsilon = 0$ is made to make the differences in the information content of the sample cases more stark, as well as to simplify computation.

exactly the same (3.5 percent). The vast majority of couples are more correctly matched under CPO (85.5 percent) than under NE (11.0 percent). We will find similar types of patterns in the complete econometric analysis reported below across different marriage market sizes both in the total U.S. sample and by geographic region and in which match-specific heterogeneity is included.

We now briefly explore another question: what is it about household pairs that make them more likely to be correctly matched under one of the two behavioral regimes? This is a more subtle question than it may first appear to be.

In our framework, the observed state variables are (w_1, w_2, Y), and given these values and the observed choice variables, we impute the parameters $(\lambda_1, \lambda_2, \delta_1, \delta_2)$ under the two different behavioral regimes. This is done for every couple in the sample, so that the likelihood of a couple being correctly matched under a given regime will not only depend on the observed state values (w_1, w_2, Y) and the imputed values of the preference and production parameters for that couple, but on the set of observed state variables and imputed values of preference and production parameters for *all* of the other couples in the sample. It is the joint distribution of all of the state variables across all households that determines the likelihood that a given couple in the sample is successfully predicted to be matched. When we assume that any household is behaving under CPO rather than NE, we know that only the imputed values of preferences, λ_1 and λ_2 are affected (since the δ_1 and δ_2 imputations are identical given our functional form assumptions). It follows that for each household in the sample, the implied values of both δ_1 and δ_2 are larger under CPO than under NE. How this affects the differentially affects the relative attraction (under CPO and NE) of a given household is difficult to determine, since the entire distribution of (λ_1, λ_2) changes under the two regimes.

The difficulty of identifying exactly what makes a couple more likely to be matched under CPO than NE is illustrated in the regression results reported in Table 2. The dependent variable in the regression is the log odds ratio for each couple of being correctly matched under the the two regimes, i.e., $\ln P(s_i^D|CPO, 2) - \ln P(s_i^D|NE, 2)$. In the panel labeled Specification 1, the dependent variable is regressed on each of the three state variables and the four endogenous time allocation values one at a time. Thus these coefficient estimates are from seven different bivariate regression specifications, and the heteroskedasticity-corrected standard errors are reported in the middle column, after the point estimates. We see from the regression results that only the time devoted to housework measures, particularly that of the males (τ_1), display any evidence that they help predict the difference in the logs of the ability to successfully match couples between CPO and NE. Couples with high levels of these time allocation measures are more likely to be better matched under the CPO regime.

In Specification 2 we estimate the multiple regression model which includes the three observed state variables and the four time allocation measures simultaneously. The findings are similar to those from the bivariate regressions, in that only the τ_1 variable has a coefficient estimate more than twice its standard error in absolute value. Thus linear regressions in which this dependent variables is considered to be an additive function of the seven observed characteristics of households have little predictive power.

One reason for the low predictive power is that the likelihood a couple being correctly matched is a function not only of their state variables, but also the entire distribution of state variables they face, as was pointed out previously. But it is also the case that the imputed values of preferences and productivities are a highly nonlinear function of wages, income, and time allocation decisions. It is likely that additive linear functions of these characteristics may do a poor job of prediction. Given the small number of observations, 282, we decided to limit attention to the three state variables that are observed, w_1, w_2, and Y, and estimated a quadratic function of these three characteristics. The results are reported under Specification 3. There are the same number of regression coefficients estimated as in Specification 2, but here we see that five of the seven coefficients have ratios of the point estimate to the standard error that are greater than 2 in absolute value. Thus taking a nonlinear function of a subset of all of the observed variables has far more ability to predict the value of the dependent variable than does a regression in which these state variables only enter linearly. Our conclusion from this exercise is that the difference in the likelihood of being correctly matched under CPO and NE may be "predictable" given couple characteristics, but only when using fairly complex transformations of those characteristics that don't lend themselves to simple interpretations.

4 Data and Empirical Results

The empirical work is performed using a sample of married couples taken from the Panel Study of Income Dynamics in the survey year 2007 that contains information on household characteristics and choices in the years 2006 and 2007. To be included in the sample, the household must have been headed by a married couple, at least one of whom was between the ages of 25 through 49, inclusive. In an effort to include individuals who had (relatively) recently been in the marriage market, we excluded all married couples who had been married for more than five years. All information on time allocations within the household must have been available for both spouses; this consists of the average amount of time spent in the labor market per week in 2006 as well as average hours spent in housework per week (reported at the time of the interview in 2007). Because household production activities change so markedly when young children are present, we excluded all households in which there was a child less than six years of age.

We also excluded any household in which one of the spouses made more than \$150 an hour or reported more than 80 hours of market work per week. We also required that the household not receive more than \$1000 per week in nonlabor income. A few households reported negative total income for the year, and these were excluded as well.

The (almost) final selection criterion imposed was that both spouses spend time both in the labor market and in home production. This, of course, is a substantive restriction that is imposed so that we can invert four first order conditions for each household to obtain four values of the unobserved characteristics of the spouses (two for each spouse).[20] Approximately 18 percent of the sample was eliminated (after imposing all other selection criteria) by insisting that both spouses report supplying time to the market in the previous year. Some spouses were reported to have supplied zero time to household production; for these individuals we assumed that the actual amount of time spent in housework was 1 hour per week.[21] During the process of estimation we found that data from 9 households in our "final" sample produced problematic values when attempting to perform the inversion required to back out unobserved preference and productivity characteristics. These sample observations were eliminated as well. The final sample included 282 married households.

It is important to point out that after imposing all of the sample restrictions, what we have is most definitely not a random sample of U.S. married households (even the full sample of the PSID cannot be considered a national random sample at this point in time). However, it is extremely difficult to say whether these restrictions tend to favor a particular test result. This is because under the two behavioral regimes, two different sets of estimates of the spouses' preference parameter are produced, and this is true for all sample members. The sorting patterns are predicted as a function of the entire joint distribution of state variables in the population under both behavioral regimes, as we discussed in Section 3.5. Because of this it is difficult to determine how sample selection affects the test results which are computed from the model of marriage market equilibrium.

The descriptive statistics and a partial summary of the first stage estimates are presented in Table 3. The descriptive sample statistics are found in the last four rows of the table. We see that husbands supply more time to the labor market than do their wives (recall that only households with both spouses working were selected), with the difference in means being about 6 hours per week. The standard deviation of market hours is larger for wives than husbands however, since a much larger proportion of women work part-time. In terms of housework, τ, average housework time was over 4 hours per week greater for

[20] If we included a household in which one member worked (say the husband) and one did not, we lose point identification of the husband- and wife-specific preferences and household productivites. In such a case, we do not observe the wage of the wife. For any possible wage that she could have received, there are a set of values of her preference parameter λ_2 which are consistent with her not working. Building a likelihood function to estimate the distribution of the state variables would involve the incorporation of conditional probabilities of sets of values of $(\lambda_2, w_2|\lambda_1, w_1, \delta_1, \delta_2, Y)$, in this particular example, and would be a daunting challenge in practice. If we were willing to make parametric assumptions on the joint distribution of the state variables, the problem becomes much more tractable. However, even in this case, recall that these distributions are required to assess the likelihood that a marriage is correctly predicted. When the choice set contains individuals with unknown values of state variables, the prediction probability requires "integrating out" the conditional distribution of the unknown state variables given the observed ones. This complication is avoided by requiring both spouses to work

[21] It would be interesting to look at the distribution of responses to these housework questions as a function of the identity of the respondent. We hazard the conjecture that, conditional on observable characteristics, respondents are likely to over-emphasize their contributions to the household workload while under-emphasizing the spouse's.

wives (10.9 versus 6.7). As has been often found, the total time spent in the labor market work and housework are quite similar for husbands and wives. We also see a substantial amount of heterogeneity in the housework time of the wives (and the husbands, though less so) as reflected in the sizeable standard deviation.

The average hourly wage paid to husbands is about 16 percent greater than the average hourly wage earned by the wives. The standard deviation of the wage distribution of the husbands is only slightly larger than the corresponding standard deviation for the wives. The similarity of the wage distributions in this case is partially attributable to the relatively few women working part-time (since a sample selection criterion was the absence of young children) and the restriction that the couple had to be married within the previous 5 years at the time of the interview. Prior to the current marriage, it is reasonable to assume that most women were working at full-time jobs, and hence may have accumulated more market human capital, resulting in higher wage offers. This rationale seems particularly germane to us, since in earlier versions of this paper with samples of couples selected without the "recent marriage" criterion, the wage distributions of the wives were decidedly less similar to the wage distributions of the husbands.

Before commenting on the nonlabor income variable we must discuss the manner in which it was constructed. For many of the components of nonlabor income, there is no way to make an attribution to a particular spouse. The fact that Y cannot be attributed uniquely to a spouse causes no problems whatsoever for the first stage estimates (where we back out the spouse-specific preference and productivity parameters) because household behavior depends solely on the sum $Y_1 + Y_2$. However, for the marriage market analysis it is necessary to have access to person-specific measures of nonlabor income. To obtain these measures, we proceeded in an admittedly arbitrary manner. We simply assumed that the each spouse's generation of nonlabor income was proportionate to their wage. Thus we had $Y_i = (w_i/(w_1 + w_2))Y$, $i = 1, 2$, where Y is total household nonlabor income (per week). If $Y = 0$, then both spouses had no nonlabor income, but if $Y > 0$, then both spouses had positive amounts of nonlabor income since $w_1 > 0$ and $w_2 > 0$ for all of our sample members. While this procedure is completely ad hoc, we don't believe our second-stage results are extremely sensitive to the imputation method. This is so because we limited all sample households to have $Y \leq 1000$, with most households having far less than this amount. Labor market earnings dwarf nonlabor income levels for the vast majority of sample members, so that the wage of a potential spouse is likely to be far more important to a potential mate than is their nonlabor income level.

The distributions of the constructed nonlabor income variables are relatively similar for husbands and wives, in large part due to the way in which they have been constructed. Husbands have higher nonlabor income, though the difference in the average amount is fairly insignificant (37.9 versus 32.9 dollars on a weekly basis). The standard deviation of the husbands' distribution was about 25 percent greater than the standard deviation of the wives' distribution.

The top first four columns of the table report the mean and standard deviations of the marginal distributions of the unobserved preference and household productivity parameters derived under the two distributional assumptions. Before discussing these estimates, note that the utility and household production function specifications deliver the implication that the first order conditions determining time spent in household production are the same under NE and CPO, and that these levels are independent of preferences. As a result, the implied values of the productivity parameters δ_1 and δ_2 are the same under NE and CPO for each household. The estimates imply that wives are more productive than husbands in housework, with the average value of δ_2 about 35 percent greater than the average value of δ_1. The amount of variation in the distributions, as summarized by the standard deviation, is approximately the same.

There is a large difference in the distributions of the preference weight on leisure for husbands and wives between the two allocation rules. Within a given regime, NE or CPO, there is not a huge difference between husbands' and wives' preferences for the private good leisure, though on average husbands value leisure more than wives. The rationale for the large differences in the preference distributions across rules is fairly intuitive. Since more of the public good is produced under the efficient allocation, two households identical in all other characteristics except preferences and mode of behavior would make the same time allocation decisions only if the efficient household valued the public good sufficiently less than did the inefficient one. This is true for every household in the sample, so that in each case the estimate of λ_1 and λ_2 must be greater under CPO than under NE. The differences are large. On average, the husbands' parameter λ_1 is 50 percent greater under CPO than under NE. For the wives, the difference is even greater, with the mean value of λ_2 under CPO being about 56 percent greater than under NE.

We now turn to describing the main focus of the empirical analysis, the likelihood-based comparisons of the predictive abilities of the two forms of household behavior in terms of marital sorting. In the results reported in Table 4, we have considered the entire U.S. as the potential marriage market for each couple in the sample (a couple from the total sample is denoted j, $j = 1, ..., 282$). As described in the previous section, we looked at potential marriage markets of size n, where $n = 2, ..., 7$. Computational limitations dictated our choice of the maximum size of marriage market be seven males and seven females.

From inspection of Table 4, we see that the likelihood of correctly predicting the marriage matches is a decreasing function of the marriage market size under both rules. This makes sense, since the choice sets are larger. We see that the log likelihood reductions are not as large at larger values of n. Not much should be read into this, since it is mainly an artifact of our assignment of a fixed value for the proportion correctly predicted for any given couple when no correct prediction occurred in any of the marriage market environments (200) in which they were embedded. As n grows large, the number of cases assigned this value increases, and the log likelihood will converge to a fixed value.[22] Our belief is

[22] Given the relatively small number of environments we are able to consider for each match, this is another reason for limiting the size of the marriage markets to be no more than seven.

that this problem should impact the log likelihood contributions under the two rules in a fairly symmetric fashion, so that we don't believe that the log likelihood ratios will be strongly affected.

We now focus attention on the last column in Table 3, which contains the log likelihood ratios for the sample. The CPO rule outperforms the NE rule for every marriage market size we have considered. In fact, the differences are increasing in the size of the labor market until we reach $n = 7$. From our discussion in the last paragraph, there is reason to believe that as n grows large the difference between the rules in predictive performance will be obscured by the inability of either to successfully predict an observed match even once over the 200 we consider. This phenomenon may well be at work in this case.

While the differences in the log likelihood ratios seem "large," we can say nothing regarding whether these are significant differences without appealing to the bootstrap results. As noted above, we drew 20 bootstrap samples for each marriage market size and computed the log likelihood ratios for each sample. Figure 2 contains the plots of the log likelihood ratios for each of the n considered. We see that for all of the statistics computed, the log likelihood ratios are positive for all n, thus indicating that the CPO intrahousehold allocation rule outperforms the NE rule in every case. Are the differences "significant?" Since $B = 20$, our nonparametric estimator of the sampling distribution of the log likelihood ratio places 0.05 mass on each value of the log likelihood ratio obtained. Thus the estimated cumulative distribution of at the minimum value of the log likelihood ratio for each n is 0.05, and since this minimum value is always positive, we can claim that these differences are "significant" at the 0.05 level. Of course, by increasing B, it seems very likely that the differences could be "significant" at substantially lower levels.

We have found what we consider to be very strong evidence that intrahousehold behavior is constrained efficient. But are potential marriage markets really national? They may be for many singles, particularly those attending college or working in a large firm, who could draw potential marriage partners from around the country of even the world. We were interested in repeating the analysis on a regional level to see how the results differed from the case of one national marriage market. We grouped couples by region of residence at the time of the 2007 PSID interview, and assumed that they lived in that region when they were in the marriage market (which was no more than 5 years prior to the interview date). For reasons of sample size, we could only use four regions: East, Midwest, South, and West. The numbers of sample households varied greatly across regions due to the cluster sampling plan of the PSID and our particular sample inclusion criteria. We have 43 sample couples from the East, 73 from the Midwest, 111 from the South, and 55 from the West.

Figure 3 contains the plots of the log likelihood ratios by marriage market size for each region. There are a total of 480 log likelihood ratios plotted in Figure 3 (20 bootstrap replications by 6 sizes by 4 regions), and we see that all but 5 are positive. Moreover, the two regions for which negative values appear are the two smallest. The largest region in terms of sample size, the South, exhibits differences that are all positive for each sample size, and the same is true of the second largest region by sample size, the Midwest. In the fourth largest sample, the East, only one negative value is observed, which was associated with $n = 6$. We could claim, in this case, that the likelihood that the true rule is CPO is 0.10 for a marriage market size of 6 in households in the East. The West exhibits 4 negative values of the log likelihood ratio, one at $n = 7$, two at $n = 4$, and one at $n = 2$. Thus we might say that the likelihood that the true rule is CPO is 0.10 for $n = 7$, 0.15 for $n = 4$, and 0.10 for $n = 2$. For all other cases in the Figure, the likelihood that NE is the rule utilized by households is no more than 0.05.

We have found what we consider to be overwhelming evidence in favor of the proposition that households behave in a constrained efficient manner. The fact that a few log likelihood ratios were found to be negative is good news in that it indicate that such a result was not preordained by the methods used to compute estimated log likelihoods. In all, including the national marriage market, we found a total of 5 negative values of the log likelihood ratio out of the 600 computed. Since we have no idea of what the actual marriage market sizes are (even given our stylized construction of them), it is reassuring to find that our inferences are essentially the same across all marriage market sizes and all regions.

5 Conclusion

In this paper we have made the point that there is no general nonparametric test available that allows one to distinguish between modes of household behavior when individual heterogeneity in unobservable and observable characteristics is not introduced in severely restrictive ways. Using a flexible specification means that within-household variation in decisions is not useful for distinguishing between competing modes of behavior, which is the negative conclusion we draw. The good news is that this heterogeneity does produce interesting implications regarding the assignments of husbands to wives in equilibrium, and that these can be exploited in investigating the mode of behavior followed by population members. Using the Gale-Shapley bilateral concept of stable matchings, we developed a new likelihood-based metric which can be used to compare the competing hypotheses of inefficient Nash equilibrium in reaction functions (NE) and efficient household behavior based on a constrained Pareto weight objective (CPO).

The methodological point we stress is reminiscent of the general problem of model overfitting. We adopted a modeling framework that was capable of perfectly fitting the data (i.e., the mapping from the data space to the parameter space was 1 to 1, otherwise known as a saturated model) under an entire class of behavioral rules $\Re$. In order to "test" one specification against another, some restrictions have to be imposed on the parameterization to make the mapping no longer 1 to 1, and to raise the possibility that one of the elements of $\Re$ "fits" better than another. Of course, the test results we obtain in the end are a function of sample realizations and the restrictions we have placed on the parametric specification of individual utilities and the household production technology. It is seldom possible to claim that one parameterization should be preferred over another on theoretical grounds alone.

Given this inherent arbitrariness, we have moved the test to a different playing field - one that is "out of sample," so to speak. The richness of the specification of individual heterogeneity leads to zero power in testing one element of $\Re$ against another using only time allocation data, but has the potential to produce the implication of different marital matching patterns - an empirical phenomenon that is not used in backing out the individual characteristics. Through an extended example, we show that the test we develop is useful for determining the mode of household behavior using the data at our disposal. Our results provide strong evidence that households do behave in a constrained efficient manner. Those advocating the "sharing rule" approach to the analysis of household allocation decisions posit efficient allocations as a fundamental identification condition. We have provided some evidence to support this assumption, though of course it remains to be shown that the constrained efficient model we utilize, CPO, actually is the best in the sense of model fit. The log likelihood function we define could, in principle, be used to consider the model fit associated with other elements of $\Re$ in addition to those to which we have limited our attention, NE and CPO.

A restrictive feature of our analysis has been the strong assumption that all (potential) marriages utilize the same rule $R \in \Re$ in making allocation decisions. In a companion paper (Del Boca and Flinn, 2012) we develop a model which allows spouses to choose between efficient and inefficient behavior, with the choice depending on all of the state variables examined here as well as a discount factor (the model is based on Folk Theorem types of arguments). It would be of great interest to allow the endogenous choice of household behavior in a marriage market context such as this, a topic we leave for future research.[23]

[23] Iyigun (2007) considers exactly this type of problem but in a strictly theoretical manner.

Table 1

Marriage Sorts from Example

Marriages		$\sigma_\varepsilon = 0.0$		$\sigma_\varepsilon = 0.1$		$\sigma_\varepsilon = 0.2$	
		NE	CPO	NE	CPO	NE	CPO
1	1						
2	2	1.000	1.000	0.927	0.862	0.640	0.587
3	3						
1	1						
2	3	0.000	0.000	0.000	0.000	0.001	0.001
3	2						
1	2						
2	1	0.000	0.000	0.071	0.136	0.275	0.328
3	3						
1	2						
2	3	0.000	0.000	0.000	0.000	0.007	0.008
3	1						
1	3						
2	1	0.000	0.000	0.000	0.000	0.005	0.007
3	2						
1	3						
2	2	0.000	0.000	0.002	0.002	0.072	0.068
3	1						

Table 2

Regression Results

Dependent Variable is $\ln L_{CPO} - \ln L_{NE}$

(s.e. estimates are heteroskedasticity-consistent)

	Specification 1			*Specification 2*		
Variable	coefficient	s.e.	ratio	coefficient	s.e.	ratio
w_1	6.838e-6	7.111e-4	0.010	-4.334e-4	8.110e-4	-0.534
w_2	1.163e-3	1.643e-3	0.708	1.637e-3	1.687e-3	0.970
Y	2.469e-5	1.110e-4	0.223	2.688e-5	9.640e-5	0.279
h_1	1.084e-3	8.085e-4	1.341	9.267e-4	8.492e-4	1.091
h_2	-5.463e-4	1.195e-3	-0.457	-6.821e-4	9.835e-4	-0.694
τ_1	5.633e-3	1.640e-3	3.435	5.647e-3	1.804e-3	3.131
τ_2	2.353e-3	1.550e-3	1.518	8.418e-4	1.622e-3	0.519

	Specification 3		
w_1	1.165e-3	1.600e-3	-0.728
w_2	-2.736e-3	1.354e-3	-2.020
Y	-1.159e-3	3.213e-4	-3.608
$w_1 * w_2$	8.695e-5	5.728e-5	1.518
$w_1 * Y$	2.653e-5	1.211e-5	2.190
$w_2 * Y$	6.847e-5	1.781e-5	3.845
$w_1 * w_2 * Y$	-1.684e-6	6.365e-7	-2.645

Table 3

Means and (Standard Deviations) of Individual Characteristics

N = 282

	Husband		*Wife*	
Characteristic	*NE*	*CPO*	*NE*	*CPO*
λ	0.401	0.603	0.372	0.579
	(0.112)	(0.103)	(0.121)	(0.124)
δ	0.070	0.070	0.095	0.095
	(0.064)	(0.064)	(0.062)	(0.062)
h	42.414		36.800	
	(10.979)		(11.620)	
τ	6.688		10.979	
	(6.215)		(8.183)	
w	21.063		18.097	
	(13.185)		(12.096)	
Y	37.893		32.895	
	(89.609)		(71.031)	

Table 4
Log Likelihood Values
U.S.
($N = 282$)

Marriage Market Size	*NE*	*CPO*	*Difference (CPO-NE)*
2	-175.448	-164.691	10.757
3	-279.686	-260.192	19.494
4	-369.825	-343.993	25.832
5	-432.488	-399.996	34.492
6	-485.183	-445.058	40.125
7	-529.887	-492.423	37.646

References

[1] Aiyagari, S. R., J. Greenwood, and N. Guner (2000) "The State of the Union" *Journal of Political Economy* 109, 213-244.

[2] Becker, G. S. (1991) *A Treatise on the Family.* Cambridge, MA: Harvard University Press.

[3] Bound, J., C. Brown, G. Duncan, and W. Rodgers (1994). "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data." *Journal of Labor Economics* 12, 345-368.

[4] Bourguignon, F., M. Browning, and P-A Chiappori. (2009) "Efficient Intra-Household Allocations and Distribution Factors: Implications and Identification," *Review of Economic Studies* 76, 503-528.

[5] Browning, M. and P-A Chiappori. (1998) "Efficient Household Allocation: A Characterization and Tests. *Econometrica* 66, 1241-1278.

[6] Browning, M., F. Bourguignon, P-A Chiappori, and V. Lachene (1994). "Incomes and Outcomes: A Structural Model of Intra-Household Allocation." *Journal of Political Economy* 102, 1067-1096.

[7] Browning, M, P-A Chiappori, and Y. Weiss (2003) "A Simple Matching Model of the Marriage Market." Manuscript, University of Chicago.

[8] Chen, Z. and F. Woolley. (2001) "A Cournot-Nash Model of Family Decision-Making." *Economic Journal* 111, 722-748.

[9] Chiappori, P-A. (1988) "Rational Household Labor Supply." *Econometrica* 56, 63-90.

[10] Chiappori, P-A. (1992) "Collective Labor Supply and Welfare." *Journal of Political Economy* 100, 437-467.

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

[12] Dagsvik, J., A. Flaatten, and H. Brunborg (2001). "A Behavioral Two-Sex Marriage Model." *Mathematical Population Studies* 9, 97-121.

[13] Del Boca, D. and C. Flinn (1994)."The Determination of Welfare in Nonintact Families." Chapter 5 in *The Measurement of Household Welfare* (Blundell, Preston, and Walker, eds.). Cambridge: Cambridge University Press.

[14] Del Boca, D. and C. Flinn (2012). "Endogenous Household Interaction." *Journal of Econometrics* 166, 49-65.

[15] Fernandez, R., N. Guner, and J. Knowles (2005). "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality." *Quarterly Journal of Economics* 120, 273-344.

[16] Gale, D. and L. Shapley (1962). "College Admissions and the Stability of Marriage." *The American Mathematical Monthly*, 69, 9-15.

[17] Greenwood, J., N. Guner, and J. Knowles (2003). "More on Marriage, Fertility, and the Distribution of Income." *International Economic Review* 44, 827-862.

[18] Iyigun, M. (2007). "Marriage Markets and Household Modes of Behavior." Working paper, University of Colorado.

[19] Iyigun, M. and R. Walsh (2007). "Building the Family Nest: A Collective Household Model with Competing Pre-Marital Investments and Spousal Matching." *Review of Economic Studies* 74, 507-535.

[20] Mazzocco, Maurizio (2007). "Household Intertemporal Behaviour: A Collective Characterization and a Test of Commitment," *Review of Economic Studies* 74, 857-895.

[21] McFadden, D. (1978). "Modelling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models* (A. Karlqvist, L. Lundqvist, F. Snikars, and J. Weibull, eds.). Amsterdam: North Holland.

[22] Pollack, R. (1990) "Two-Sex Demographic Models" *Journal of Political Economy* 98, 399-420.

[23] Roth, A. and M. Sotomayor (1990). *Two-sided Matching; A Study in Game-theoretic Modeling and Analysis*. Cambridge: Cambridge University Press.

[24] Seitz, S. (1999) "Employment and the Sex Ratio in a Two-Sided Model of Marriage." Manuscript, Queens University.

Figure 1
Difference in Likelihood by Couple
CPO minus NE

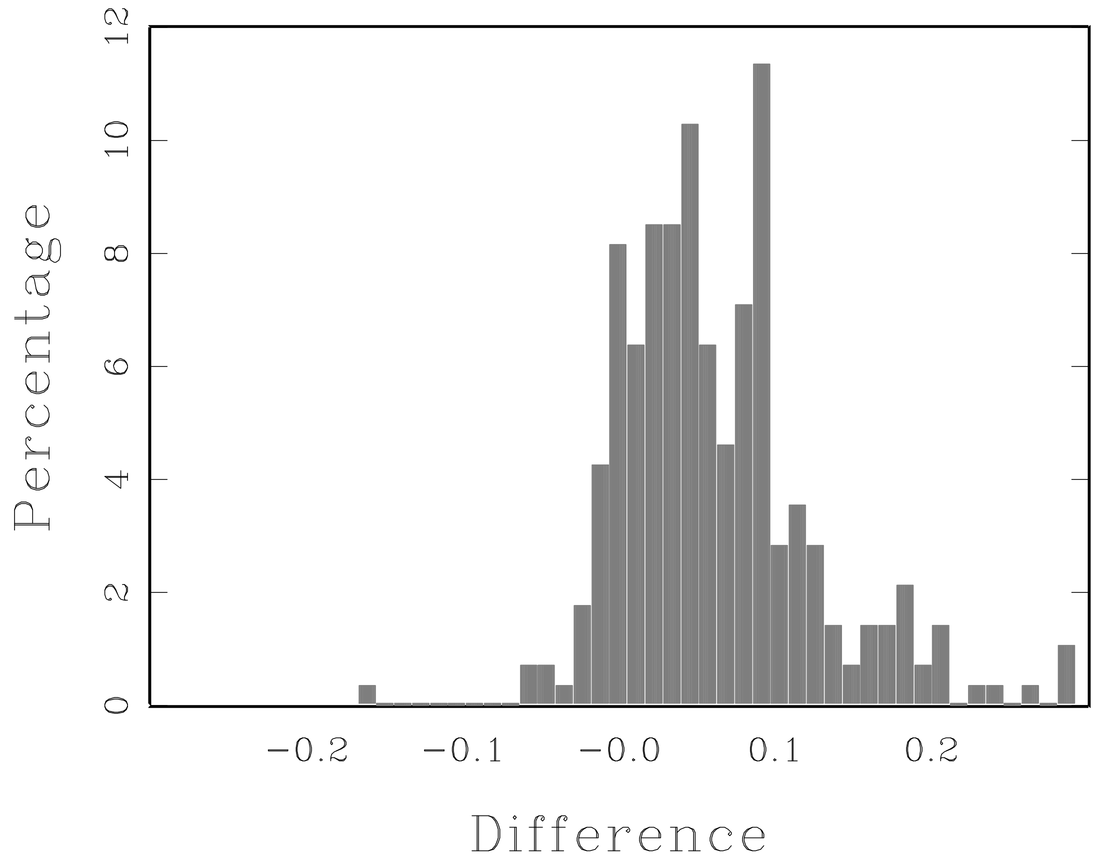

Figure 2
Ln Likelihood Difference
Entire U.S.

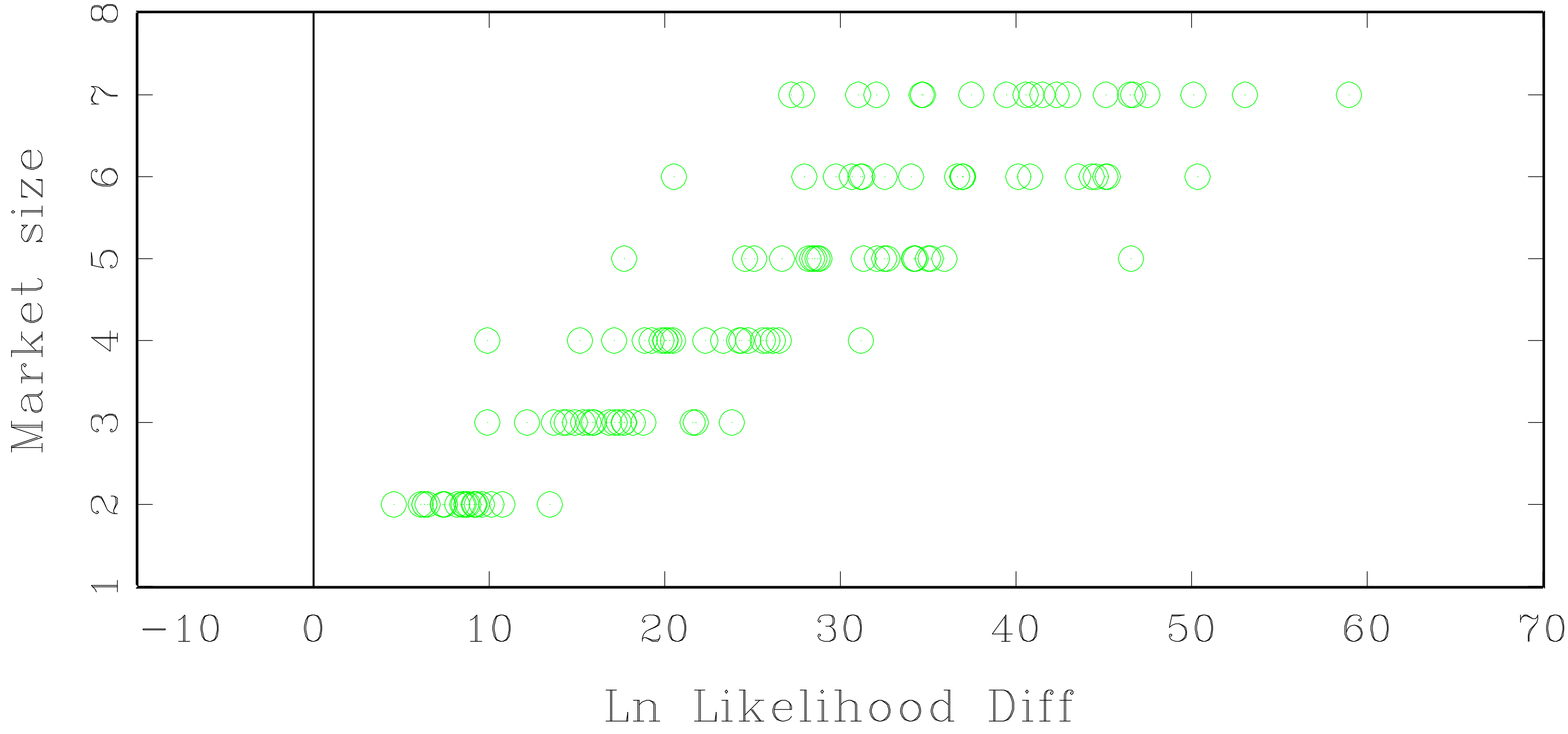

Figure 3.a
Ln Likelihood Difference
East

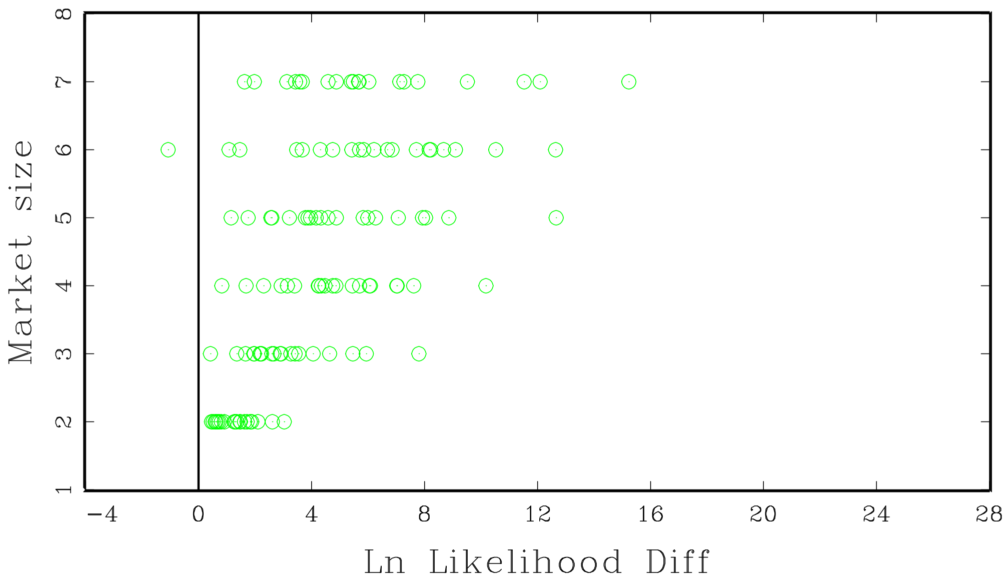

Figure 3.b
Ln Likelihood Difference
Midwest

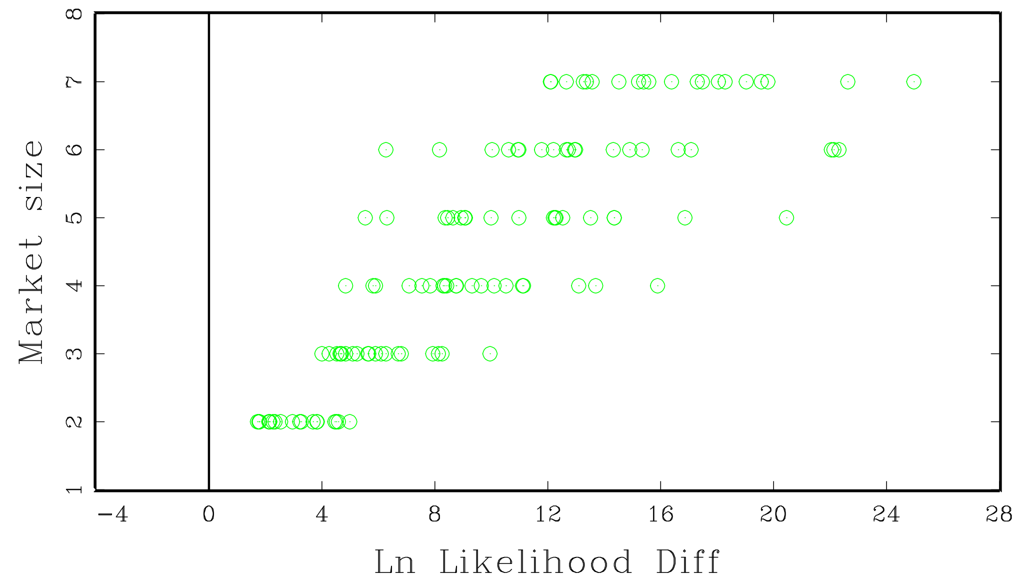

Figure 3.c
Ln Likelihood Difference
South

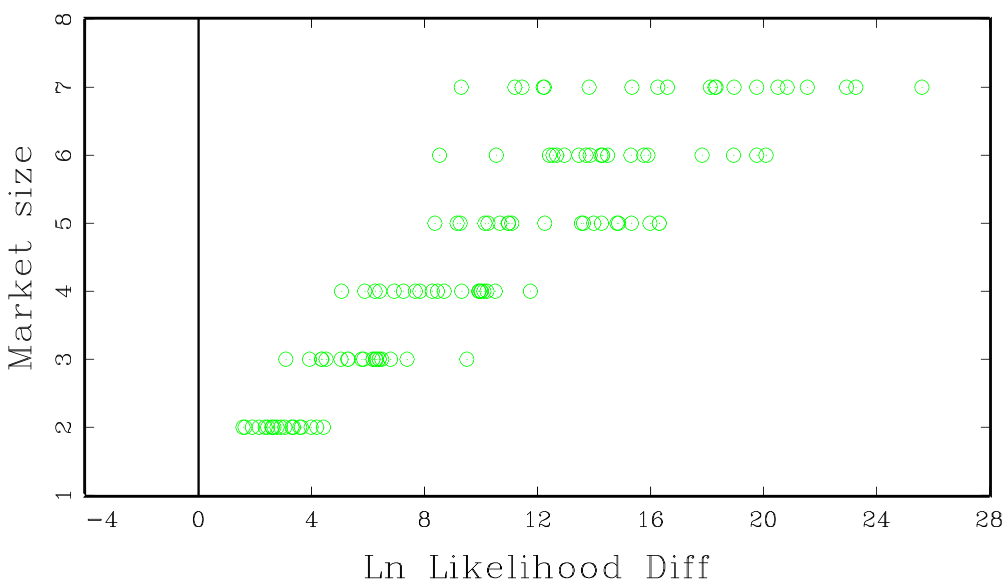

Figure 3.d
Ln Likelihood Difference
West