Lifecycle marriage matching: Theory and evidence*

Eugene Choo and Aloysius Siow
University of Toronto
September 29 2006

Abstract

The marriage market is a dynamic bilateral matching market with finitely lived participants. The growth rate of the number of marriages between two types of individuals estimates the systematic benefit to them delaying marriage for a period versus marrying immediately. Identification of the semi-parametric model is transparent. The model is estimated with 1990 US marriage data. The estimated model shows that a concern for accumulating marriage specific capital is quantitatively significant in generating positive assortative matching in spousal ages at marriage, gender differences in spousal ages at marriage, and a preference for early marriage. Gender variations in population supplies due to gender specific mortality rates and entry cohort sizes have offsetting quantitative effects.

1 Introduction

The bivariate distribution of spousal ages at marriage (hereafter marriage distribution) has some well known empirical regularities (Casterline, et. al. 1986). First, there is positive assortative matching in spousal ages at marriage. Second, except for the youngest age groups, marriage rates decline with age. Third, men generally marry younger women. These regularities are shown in Figures 1a), b) and c) which plots the marriage distribution for 1990 for the US. To the best of our knowledge, no existing theory explains all three regularities.[1]

We propose a general equilibrium behavioral framework for estimating and describing marriage distributions. It captures many features of the marriage distribution. The framework also allows us to identify marital preference parameters using a system of differential equations. Our identification strategy of these primitives is transparent, straightforward and imposes little apriori restrictions on the data. The empirical model also allows us to embed and test different theories of marital or household production. We quantitatively explore the efficacy of different theories in explaining the different features of the marriage distribution.

*We thank Angelo Melino and Joanne Roberts for comments and insights, and Shashi Khatri for excellent research assistance. We also thank seminar participants at ESPE, NEIPA, Johns Hopkins University, Northwestern University, University of British Columbia, UC Berkeley, University of Toronto, University of Wisconsin for useful comments. Choo and Siow both thank SSHRC for financial support.

[1] Partial explanations include Anderberg 2003; Becker 1973, 1974, 1991; Bergstrom and Bagnoli 1993; Bergstrom and Lam 1989; Gould and Passerman 2003; Siow 1998. Waite, et. al. 2000 include discussions of models by non-economists.

Like the growing body of empirical literature on marital matching, our central focus is better understanding of preferences underlying marital matching. Consider an environment where we observe data of final matches between heterogenous types of males and females, together with the respective supplies of available individuals. We further assume that these equilibrium matches are drawn from a stationary data generating process where preferences are invariant over time. Given these restrictions on what we observe, we propose an empirical model that allows us to estimate preferences consistent with the pattern of aggregate matches. Then we ask to what extent do these preferences affect the marital sorting that we obseve.

Our behavioral empirical framework is based on the following observation:

The growth rate of the number of marriages between two types of individuals approximates the systematic benefit to them delaying marriage for one period versus marrying immediately. This systematic benefit is independent of marriage market conditions.

This result is intuitive. If there is an increase in the number of marriages in the next period between two types of individuals relative to today, the systematic benefits from delaying marriage for one period must have increased. The growth rate interpretation is motivated by the arbitrage argument familiar from intertemporal choice models in competitive environments.

Our equilibrium behavioral model provides a behavioral interpretation to the growth rates of marriages. This opens the door for behaviorally motivated parameterizations of marital output. Our model suggests that growth rates identify preference parameters that are independent of marriage market supply conditions. Independent of the behavioral interpretation, this paper suggests that the growth rates of marital matches and associated terminal values are useful summary statistics for marriage distributions. We show that the marriage distribution can be modelled using a system of differential or growth rate equations. Together with a set of terminal conditions and accounting equations describing how these aggregate quantities evolve, this system of differential equations can characterize any observed marriage distribution.

Becker (1973, 1974) proposed the first classic explanation for positive assortative matching in static transferable utility models.[2] If spousal attributes are vertically differentiated where more of a spousal attribute is desirable, Becker showed that positive assortative matching can be explained assuming complementarity in marital output by spousal attributes and marriage market clearing. Becker's insight is not however not directly applicable to many horizontal attributes like race, religion and age. Furthermore, attributes like age of an individual changes over time. Instead of complementarity in spousal attributes in marital output production, we propose a decomposition of marital output that comprise only of separable spousal contribution and marriage specific capital.[3] We estimate and test the efficacy of this marital output decomposition in explaining the observed pattern of marriage matching. Our estimation methodology does not impose any apriori own age effects on marital output. Nor does our empirical model contain any specific behavioral gender differences. Rather, we fit gender specific parameters in the model which are consistent with observed gender differences in the data.

More formally, let π_{ij}^k, denote the marital output in the k'th year of marriage for a couple who marry when the male and female are of age i and j respectively. We consider a decomposition of marital

[2]Empirical development of his model include Grossbard-Shectman 1993; Suen and Lui 1999.

[3]We do not allow any complimentarities in the form of spousal age interaction effect on marital output.

output into separable spousal contributions and marriage specific capital without complementarities as suggested by Becker. We want to explore the efficacy of this parameterization in explaining the marriage distribution. The decomposition we have in mind takes the form,

$$\pi_{ij}^{k} = \alpha_{i+k} + \gamma_{j+k} + \kappa_k + \nu_{ijk}.$$

α_{i+k} and γ_{j+k} are the husband's and wife's contributions respectively. κ_k represents the value of current marital output that is generated by the stock of marriage specific capital. This stock, κ_k, depends on how long the marriage has lasted. In other words, marital output depends on what each spouse brings to the table and how long the couple has been together. Let μ_{ij} denote the number of observed matches between an age i males and an age j female. Our empirical model shows that the growth rate of marriage in the form of $\ln(\mu_{i\,j}/\mu_{i+1\,j+1})$ approximates the invariant marital gain to an (i, j) couple marrying today as opposed to delaying for one period. Imposing separable spousal contribution with marriage specific capital on marital output, we will show that the growth rate in marriage can be reparameterize as

$$\ln \frac{\mu_{ij}}{\mu_{i+1,j+1}} = \frac{1}{2}(\alpha_i - \alpha_{i0}) + \frac{1}{2}(\gamma_j - \gamma_{0j}) + \frac{1}{2}\kappa_{ij} + \psi_{ij}. \tag{1}$$

The preference parameters of interest are $\frac{1}{2}(\alpha_i - \alpha_{i0})$ and $\frac{1}{2}(\gamma_j - \gamma_{0j})$, which are the male and female marital contributions relative to their mean utilities from being single respectively, and $\frac{1}{2}\kappa_{ij}$ which captures the maximum level of marriage specific capital that an (i, j) couple can attain together. ψ_{ij} is a composite error term. The form of equation 1 also allows us to identify these preference parameters using a simple dummy variable regression.

We estimate our model using marriage data from the U.S. for 1990. The estimated models were used to simulate marriage distributions to investigate the quantitative importance of different determinants of the marriage distribution. Summarizing our empirical investigation, gender specific spousal contributions and marriage specific capital are important determinants of the marriage distribution. A preference for an own age spouse is quantitatively unimportant. Our model estimates suggest that gender differences in mortality and birth rates can explain around 0.4 years of the average age gap between men and women. While supply factors are significant, gender specific mortalities are quantitatively more important than gender differences in initial supplies.

In terms of overall fit, our simple stationary model can replicate well know features of the marriage distribution. It can account for the large marriage rates of young adults, positive assortative matching in spousal ages of marriage, the variations in the ages of marriage and gender difference in spousal ages. Perhaps the largest discrepancy is that we over predict the average ages of marriage by ten percent. We also under predict the number of middle aged unmarrieds and over predict the number of older unmarrieds. We show that these last discrepancies are due primarily to our stationarity assumptions.

Our estimates suggest that specific marriage contributions are surprisingly similar across gender. We also find that marital specific capital is quantitatively important in rationalizing regularities of the marriage distribution like positive assortative matching in marital ages.[4] Since individuals have finite lives and the accumulation of marriage specific capital requires time, a young individual who marries an older spouse will anticipate having less time to accumulate and enjoy marriage specific capital relative

[4]Examples of marriage specific capital include investments in durable assets like housing, retirement savings, or children. These within marriage investments lose value or are captured by one spouse when the marriage ends (Stevenson 2005).

to marrying a same age or younger spouse. From this perspective, younger spouses prefer to marry each other. Older individuals do not value younger spouses more than older spouses since they have shorter expected lifespans, and so will be willing to marry each other. Marriage specific capital also generates a demand for marrying early. An early marriage allows the couple to accumulate more marriage specific capital. Thus marriage specific capital can explain why marriage rates in general decline with age.

We also ask the question, what would the marriage distribution have been like if there were no gender differences in both mortality and birth rates. What would be the extent of assortative matching without these gender differences? The birth rates in the US for males are slightly higher on average than for females.[5] However, mortality hazard increases much faster with age for men. Even with more boys than girls being born, women still outnumber men in the United States. In 2003, the Census Bureau estimated a total of 144,513,361 females of all ages, compared to 138,396,524 males. As women have longer life expectancies than men, older men are scarce relative to older women in the marriage market. This relative scarcity might advantage older men in the marriage market providing them with better marriage possibilities with younger women.

This paper builds on a large literature. Methodologically, our formulation of dynamic marriage matching borrows heavily from the large body of literature on dynamic discrete choice models. We adopt the conditional independence and additive separability assumptions first proposed by Rust (1987). We employ the invertibility result of Hotz and Miller (1993) when representing the value functions in terms of choice probabilities. Aguirregabiria and Mira (2002) proposed an important alternative representation of the solution to the dynamic programming problem in the choice probabilities space.[6] This alternate mapping is the building block of their nested pseudo-likelihood estimation procedure which nest the likelihood based estimation approach proposed by Rust(1987) and Hotz and Miller(1993).

We diverge from these works in the way we estimate our model. To minimize computational cost, we proposed a one to one transformations of observed choice probabilities that is consistent with our behavioural general equilibrium model. These transformations of choice probabilities are economically meaningful and have behavioural interpretations. In our Walrasian equilibrium, match specific transfers equate the demand and supply of spouses which can be thought in terms of transformed choice probabilities. The resulting equilibrium restriction implied by our model permits the straightforward estimation that we adopt. Our estimation strategy is closes in spirit to the conditional choice estimator probabilities estimator of Hotz and Miller (1993). This point of departure from existing empirical dynamic discrete choice literature will be made clear in the model section of this paper. Our estimation strategy of per period match utility mirrors the exact identification results of per period profits proposed by Pesendorfer and Schmidt-Dengler (2003).

Most closely related to our work are the papers in the empirical search/matching literature, in particular Wong(2003) and, Brien, Lillard and Stern (2006). Like ours, these paper uses data on final match ourcomes and the attributes of individuals in these matches. The identification of preferences relies on structural assumptions about the marriage market mechanism by which these final matches are achieved. We all impose equilibrium restrictions in our estimation procedure. Brien, Lillard and

[5] For example, according to the National Vital Statistics Report, 2005, the highest sex birth ratio occurred in 1946 (1,059 male births per 1,000 females) while the lowest occurred in 1991 and again in 2001 (1,046 male births per 1,000 females).

[6] They show that the choice probabilities that solves the dynamic programme is the unique fixed point of this alternate mapping in probability space.

Stern (2006) which looks the evolution of match value to better understand cohabitation and divorce is also much more ambitious in scope than our work. The main difference between our work and that of Wong(2003) and, Brien, Lillard and Stern (2006) are the assumptions made about the marriage market mechanism. Instead of using search frictions to model differences in delay of marriage for ex-ante identical individuals, we assume that delay differences are due to idiosyncratic realizations of payoffs to delay and marriage. Our random utility approximations of the marriage market mechanism allows us to derive analytical equilibrium conditions on aggregate matches. This in turn afford us with a computationally simple estimation procedure. We are intellectually indebted to search models which first modelled delay differences as optimal responses to differences in per period idiosyncratic payoffs from different actions. Our dynamic transferable utility framework also uses match specific transfers to clear the marriage market.[7] This paper, to our knowledge, is the first to estimate the marriage matching distribution (bivariate distribution of who matches with whom by ages of entry into the market) of a stationary dynamic bilateral matching model.

Our work is also closely related to Hitsch, Hortaçsu and Ariely (2005). The authors uses a detailed search information from a novel dataset of online dating to infer revealed spousal preference. The detailed nature of the dataset allows the researchers to observe actual choices as well as the choice sets available to participants. This enviroment also allows them to estimate a rich specification over spousal physical and socio-economic characteristics using a straightforward estimation strategy. Unlike our approach (and that of Wong(2003) and Brien, Lillard and Stern (2006)), their estimation procedure does not rely on assumptions on how final matches arise. On the other hand, their estimation approach ignores any strategic behavior in matching. The empirical prediction of these preference estimates were then investigated using simulations from the Gale-Shapley model.

There is also an active literature on marital matching in dimensions other than spousal ages. E.g. Anderberg, Anderson 2003; Bisin, et. al. 2004; Boulier and Rosensweig 1984; Chiappori, et. al. 2005, Chiappori and Reny 2005; Choo Siow forthcoming; Compton and Pollak 2004; Costa and Kahn 2000; Guner, et. al. 2004, Fogli, et. al 2004, Iyigun and Walsh 2004; Lam 1988; Lich-Tyler 2003; Lundberg and Pollak 2003; Mare 1991; Peters and Siow 2002; Qian and Preston 1993. Our model is complementary to that literature. Fox (2005) proposes a general maximum score estimator for the match production function in an endogenous transferable utility framework. This estimator implicitly embeds inequalities implied by single-agent best responses. Bajari and Fox (2005) applies this approach to estimate the payoffs of bidders in multiple-unit spectrum auction. Bergstrom and Lam (1989) have the first quantitative dynamic transferable utility model of the marriage market where individuals care about when they marry, but not to whom they marry. They use this model to study how the marriage market adapts to demographic shocks (Also see Foster and Khan 2000).

2 The Model

In this section, we outline the setup of our dynamic full-commitment model of marriage with exogenous divorce. We will focus on the central features of our model. We will also discuss the models' equilibrium

[7] Mortensen 1988 introduced search models to explain delay in marriage among otherwise observationally equivalent individuals. Also see Anderson and Smith 2004; Burdett and Coles 1997, Shimer and Smith 2000. Empirical applications include Ayagari, et. al. 2000, Anderberg; Hamilton and Siow 2002, Seitz 2002. Wong (2005) proposes to estimate a search model whose aim is also to rationalize the marriage distribution.

conditions that we take to data. Our objective is identify the marital payoff parameters that are implied by these equilibrium conditions. We use these parameters to test various theories rationalizing the marriage distribution. In this dynamic generalization, we also adopt the approximation of the marriage market first introduced in Choo and Siow (2006). While this framework makes simplifying assumptions about how heterogenous individuals match in the marriage market, it provides tractable equilibrium conditions that can easily be taken to data. We will show that the model captures much of the features of the marriage distribution despite these simplifying assumptions about the marriage market. Our framework also employs the dynamic discrete choice framework introduced by Rust(1987).

Consider a stationary society populated by overlapping generations of adults. For expositional convenience, we assume that each individual lives for Z periods. We relax this assumption in the empirical work by allowing for differential mortality by age and gender. The youngest adult is of age one. The age of a male is indexed by i and the age of a female is indexed by j. In any period, the type of an adult is defined by his or her age. We focus on empirically characterizing the marriage distribution by age. While we acknowledge that age is not a complete characterization of type, we believe that our framework fills a gap in the empirical marriage matching literature. Our proposed methodology can be easily extended to deal with fixed horizontal attributes or types like race and religion.[8]

The society is stationary in the sense that the same numbers of age one males and age one females, m_1 and f_1 respectively, enter the society at the beginning of each period. Let m_i and f_j denote the endogenous numbers of single males of age i and females of age j respectively who are looking for spouses at the beginning of each period.

Single adults can choose to remain unmarried or marry. At each period a single individual faces a random utility draw from each type of spouse available and from remaining single that period. He or she chooses the option that maximizes his or her discounted expected utility. In the case of a single male, if he decides to marry a particular type of female, he commits to pay a once off transfer specific to these two types of individuals matching. Similarly, single females who decide to marry a particular type of men agree to receive this equilibrium transfer that clears the marriage market. So if individual g of type i wants to marry a woman G of type j, he has to transfer τ_{ij} amount of marital output to her. Similarly, if woman G of type j wants to marry man g of type i, she has to be willing to accept τ_{ij} amount of marital output from him. Each individual takes τ_{ij} as exogenous. The marriage market clears when given τ_{ij} for every i, j, the number of type i men who want to marry type j women is equal to the number of type j women who want to marry type i men. In this full-commitment model, the one time payment of τ_{ij} fully internalizes the discounted stream of within marriage utilities for this couple, the exogenous divorce probabilities and the relative scarcity of males and females in the system.

Denote the first year of marriage as year 0. The time discount factor is denoted by β. Marriages may end in divorce or the death of a spouse. Divorce occurs at some exogenous rate, δ in which case divorced individuals re-enter the marriage market. Let the survival probability of marriage be denoted by $S = 1 - \delta$. If he or she decides to remain single, he or she receives some random utility this period and re-enters the marriage market next period as an older man or woman. Single adults who do not marry, divorced adults, and adults whose spouses die will re-enter the marriage market in the next period. We do not distinguish available adults by previous marital status.

We assume that the divorce hazard does not vary by type of couples nor duration of marriage. This

[8]These attributes are either fixed or evolves deterministically over time.

approximation though restrictive simplifies the equilibrium condition which we take to data and our estimation strategy. In an earlier draft, we allowed the exogenous divorce hazard to differ by age of marriage and the tenure of marriage with little additional empirical payoff. Section 8.1 in the Appendix provides a discussion of the model and alternate estimation approach when we do not make this approximation. We will show that despite this approximation, our model does a reasonable job at capturing the marriage distribution.

Male Decision Problem:

Formally, a single male individual g in each period is characterized by two state variables. His age denoted by i, and a $(Z+1)$ vector of i.i.d idiosyncratic payoffs or match specific errors, $\boldsymbol{\varepsilon}_{ig}$, that is unobserved to the econometrician. The i.i.d. vector of random variables, $\boldsymbol{\varepsilon}_{ig}$ us drawn from MacFadden's type I extreme value distribution, denoted by $F(\cdot)$. Let a denote the decision of this individual in each period where $a \in \mathcal{D} = \{0, 1, \ldots, Z\}$, and corresponds to the decision to not marry $a = 0$, and to marry a spouse of age j, $a = j$. $v_i(a, \boldsymbol{\varepsilon}_{ig}, \boldsymbol{pi})$, is the discounted sum of current and future payoffs from action a where $\boldsymbol{\pi}$ denotes a vector of parameters defining preferences. Naturally, he chooses the payoff maximizing action. The value function for this single individual is given by

$$\begin{aligned} V(i, \boldsymbol{\varepsilon}_{ig}, \boldsymbol{pi}) &= \max_{a \in \mathcal{D}} \; v_i(a, \boldsymbol{\varepsilon}_{ig}, \boldsymbol{pi}), \\ &= \max_{a \in \mathcal{D}} \; \big\{ v_i(0, \boldsymbol{\varepsilon}_{ig}, \boldsymbol{pi}),\; v_i(1, \boldsymbol{\varepsilon}_{ig}, \boldsymbol{pi}), \ldots,\; v_i(Z, \boldsymbol{\varepsilon}_{ig}, \boldsymbol{pi}) \big\}. \end{aligned} \tag{2}$$

Our specification of preferences over partners satisfy the additive separability and conditional independence assumptions introduced by Rust (1987). If he chooses to remain single, that is, $a = 0$, his payoff from this decision takes the form,

$$v_i(0, \boldsymbol{\varepsilon}_{ig}, \boldsymbol{pi}) = \alpha_{i0} + \varepsilon_{i0g} + \beta \boldsymbol{V}_{i+1}, \tag{3}$$

where $\boldsymbol{V}_{i+1}$ denotes the the expected value function (or the integrated value function), that is,

$$\boldsymbol{V}_{i+1} = \mathbb{E} V(i+1, \boldsymbol{\varepsilon}_{i+1\,g}, \boldsymbol{pi}) = \int V(i+1, \boldsymbol{\varepsilon}_{i+1\,g}, \boldsymbol{pi}) \; dF(\boldsymbol{\varepsilon}_{i+1\,g}).$$

where the expectation operator, $\mathbb{E}$, is taken with respect to the idiosyncratic vector of match errors, $\boldsymbol{\varepsilon}_{i+1\,g}$. In the current period he receives a payoff of α_{i0}, which is common to all age i individuals, and an individual specific i.i.d. idiosyncratic payoff of ε_{i0g}. The term $\beta \boldsymbol{V}_{i+1}$ is the discounted expected payoff from participating in the marriage market next period when he is one year older.

Suppose he chooses to marry an age j female who is older than himself, that is $a = j > 0$, where $j > i$, then his payoff takes the form,

$$v_i(j, \boldsymbol{\varepsilon}_{ig}, \boldsymbol{pi}) = \frac{1}{2} \sum_{k=0}^{z_{ij}} (\beta S)^k \, \pi_{ij}^k - \tau_{ia} + \varepsilon_{ijg} + \beta\delta \sum_{k=0}^{z_{ij}-1} (\beta S)^k \boldsymbol{V}_{i+k+1} + \beta^{z_{ij}+1} S^{z_{ij}} \boldsymbol{V}_{i+z_{ij}+1}.$$

π_{ij}^k be the $k'th$ period marital output of a couple who married when the male was age i and the female was age j. π_{ij}^k is realized only if the marriage did not result in a divorce or death of a spouse previously. Let $z_{ij} = \min(Z - i, Z - j)$, where $z_{ij} + 1$ measures the maximum number of years that the marriage can attain. The first term $\sum_{k=0}^{z_{ij}} (\beta S)^k \, \pi_{ij}^k$ is the present discounted value of within marriage output in the event that the marriage does not dissolve. This is shared equally among the spouses. The second term τ_{ij} denotes the one time transfer the husband makes to his wife. This transfer is common to all marriages between i type man and type j woman and it can be positive or negative. The term ε_{ijg} is the current period individual specific idiosyncratic payoffs g receives from matching with a type j female. Marriages dissolve at an exogenous rate δ each period. In each period of this marriage (other than the last period, z_{ij}), g has a probability δ of reentering the marriage market. The term $\sum_{k=0}^{z_{ij}-1} \beta^{k+1} S^k \delta \boldsymbol{V}_{i+k+1}$ represents this expected discounted value of reentering the marriage market as the individual ages. Since $j > i$, g reenters the marriage market with certainty after the last period of marriage. This expected discounted value is represented by $\beta^{z_{ij}+1} S^{z_{ij}} \boldsymbol{V}_{i+z_{ij}+1}$. ε_{ijg} is the realization of an i.i.d. random variable with type I extreme value distribution.[9]

The distributional assumption on the idiosyncratic match error $\boldsymbol{\varepsilon}$ allows the equilibrium probability that a type i male matches with a type j female to have the familiar multinomial logit form. Let $\mathcal{P}_{i0} = \mathcal{P}_i(j = 0)$ be the equilibrium probability that an i type male stays single and $\mathcal{P}_{ij} = \mathcal{P}_i(j)$ be the probability that i marries a type j female. $\mathcal{P}_{ij}$ with the following well know multinomial form,

$$\begin{aligned} \mathcal{P}_{i0} &= \int \{0 = \arg\max_{a \in \mathcal{D}} \; v_i(a, \boldsymbol{\varepsilon}_{ig}, \boldsymbol{\pi}) f_{\varepsilon}(\boldsymbol{\varepsilon}_{ig}) d\boldsymbol{\varepsilon}_{ig} \\ &= \frac{\exp(\alpha_{i0} + \beta \boldsymbol{V}_{i+1})}{\exp(\alpha_{i0} + \beta \boldsymbol{V}_{i+1}) + \sum_{r=1}^{Z} \exp(\, \frac{1}{2} \sum_{k=0}^{z_{ir}} (\beta S)^k \, \pi_{ir}^k - \tau_{ir} + \boldsymbol{A}_{ir} \,)}, \end{aligned} \tag{4}$$

where $\boldsymbol{A}_{ir} = \beta\big[\delta \sum_{k=0}^{z_{ir}-1} (\beta S)^k \boldsymbol{V}_{i+k+1} + (\beta S)^{z_{ir}} \boldsymbol{V}_{i+z_{ir}+1}\big]$.[10] The Bellman equation also has the following familiar form,

$$\begin{aligned} \boldsymbol{V}_i &= \log \Big\{ \exp\big(\alpha_{i0} + \beta \boldsymbol{V}_{i+1}\big) + \sum_{r=1}^{Z} \exp\big(\, \frac{1}{2} \sum_{k=0}^{z_{ir}} (\beta S)^k \, \pi_{ir}^k - \tau_{ir} + \boldsymbol{A}_{ir} \,\big) \Big\} + c \\ &= \alpha_{i0} + c + \beta \boldsymbol{V}_{i+1} - \log \mathcal{P}_{i0}. \end{aligned} \tag{5}$$

The parameter c is the Euler's constant which arised from our distributional assumption on the unobserved idiosyncratic deviations in marital output. Equation 5 says that the value of participating in the marriage market at age i denoted by $\boldsymbol{V}_i$ can be divided into two components. The first is the expected utility from being single this period which comprise of $\alpha_{i0} + c$ and the expected value of participating in the marriage market next period as represented by $\beta \boldsymbol{V}_{i+1}$. The second term, $-\log \mathcal{P}_{i0}$ captures the expected utility from choosing to be married at age i. If the marriage rate for type i males is high, (or the probability of being single, $\mathcal{P}_{i0}$ is low), then the expected utility from being married is high.

Taking the log-odds ratio of two alternatives and repeated substitution of 5 and 4, we get

$$\log \Big\{ \frac{\mathcal{P}_{ij}}{\mathcal{P}_{i0}} \Big\} = \sum_{k=0}^{z_{ij}} (\beta S)^k \Big(\frac{1}{2} \pi_{ij}^k - \alpha_{i+k,0} \Big) - \sum_{k=1}^{z_{ij}} (\beta S)^k \Big(c + \ln \mathcal{P}_{i+k,0}^{-1} \Big) - \tau_{ij}. \tag{6}$$

Equation 6 gives the difference in systematic expected payoffs for an i type male marrying a j type female relative to remaining single that period. The term $\sum_{k=0}^{z_{ij}} (\beta S)^k \, \pi_{ij}^k$ represents the stream of expected utility he gets from the match in the event that the marriage does not dissolve. At each age $i + k$

[9] In the case when $j \leq i$, the payoff from marriage has the form

$$v_i(j, \boldsymbol{\varepsilon}_{ig}, \boldsymbol{\pi}) = \frac{1}{2} \sum_{k=0}^{z_{ij}} (\beta S)^k \, \pi_{ij}^k - \tau_{ia} + \varepsilon_{ijg} + \beta \sum_{k=0}^{z_{ij}-1} (\beta S)^k \delta \boldsymbol{V}_{i+k+1}$$

[10] Again, this applies to the case where $j > i$. When $i \geq j$, then $A(i, r) = \beta\delta \sum_{k=0}^{z_{ir}-1} (\beta S)^k V_{i+k+1}$.

of marriage, the difference in expected utility from being locked in marriage and participating in the marriage market that period is represented by $\alpha_{i+k,0} - \beta S(c + \ln \mathcal{P}^{-1}_{i+k+1,0})$. Again $\ln \mathcal{P}_{i+k,0}$ is a statistic for the gains from the marriage market at $i+k$. In choosing this match, the i type male commits to pay a match specific transfer τ_{ij} to his spouse.

Female Decision Problem:
The specification of the decision problem of a woman is similar. The set of marital choices available to a single female at any point in time is given by $\mathcal{D} = \{0, 1, \ldots, Z\}$, that is the set of male ages and remaining unmarried. The payoffs that a female individual h of age j receives from pursuing action a is denoted by $w_j(a, \boldsymbol{\varepsilon}_{jh}, \boldsymbol{\pi})$ The value function of a single type j female is denoted by,

$$W(j, \boldsymbol{\varepsilon}_{jh}, \boldsymbol{\pi}) = \max_{a \in \mathcal{D}} \; w_j(a, \boldsymbol{\varepsilon}_{jh}, \boldsymbol{\pi}), \tag{7}$$

and the expected value before she sees her idiosyncratic state vector, $\boldsymbol{\varepsilon}_{jh}$ is given by,

$$\boldsymbol{W}_j = \mathbf{E} W(j, \boldsymbol{\varepsilon}_{jh}, \boldsymbol{\pi}) = \int W(j, \boldsymbol{\varepsilon}_{jh}, \boldsymbol{\pi}) \; dF(\boldsymbol{\varepsilon}_{jh}).$$

If her payoff maximizing choice was to marry a type i male, where $i > j$, her discounted payoff is given by,

$$w_j(i, \boldsymbol{\varepsilon}_{jh}, \boldsymbol{\pi}) = \frac{1}{2} \sum_{k=0}^{z_{ij}} (\beta S)^k \, \pi^k_{ij} + \tau_{ij} + \varepsilon_{ijh} + \beta\delta \sum_{k=0}^{z_{ij}-1} (\beta S)^k \boldsymbol{W}_{j+k+1} + \beta^{z_{ij}+1} S^{z_{ij}} \boldsymbol{W}_{j+z_{ij}+1}. \tag{8}$$

The interpretation of the parameters is similar to the male counterpart. Recall that π^k_{ij} is the marital output in the k'th period of marriage for a couple that marry when the male and female ages are i and j respectively. The first term represents the female share of the discounted within marriage payoffs when the marriage does not dissolve. If she agrees to this match, she receives an equilibrium transfer τ_{ij} from her partner. These first two terms are common to all (i, j) matches. The term ε_{ijh} is the idiosyncratic payoffs specific to her random match draw. Given that divorce occurs at an exogenous rate δ, the term $\beta\delta \sum_{k=0}^{z_{ij}-1} (\beta S)^k \boldsymbol{W}_{j+k}$ captures the expected value of re-entering the marriage market in the future in the event of divorce. If she marries an older man $(i > j)$ and her marriage does not dissolve, she reenters the marriage market with certainty after the last period of the match. The term $\beta^{z_{ij}+1} S^{z_{ij}} \boldsymbol{W}_{j+z_{ij}+1}$ represents the discounted expected value of being single at age $j + z_{ij} + 1$.[11] If she chooses to remain single, that is $a = 0$, then

$$w_j(0, \boldsymbol{\varepsilon}_{jk}, \boldsymbol{\pi}) = \gamma_{0j} + \varepsilon_{0jh}, + \beta \boldsymbol{W}_{j+1}.$$

The first term γ_{0j} is the systematic payoff to j from remaining single which is common to all j females. The interpretation of the remaining parameters is analogous to the male counterpart of equation (3).

Let $\mathcal{Q}_{ij}$ and $\mathcal{Q}_{0j}$ denote the probability that a j type female marries an i type male and remains single respectively. The log-odds ratio that a j type female marries an i type male relative to remaining

[11] If she marries someone younger, i.e. $i < j$, the payoff from this choice would not have the last term of equation (8). That is,

$$w_j(i, \boldsymbol{\varepsilon}_{jh}, \boldsymbol{\pi}) = \frac{1}{2} \sum_{k=0}^{z_{ij}} (\beta S)^k \, \pi^k_{ij} + \tau_{ij} + \varepsilon_{ijh} + \beta\delta \sum_{k=0}^{z_{ij}-1} (\beta S)^k \boldsymbol{W}_{j+k+1}$$

single equals

$$\log\left\{\frac{\mathcal{Q}_{ij}}{\mathcal{Q}_{0j}}\right\} = \sum_{k=0}^{z_{ij}}(\beta S)^k\left(\frac{1}{2}\pi_{ij}^k - \gamma_{0j+k}\right) - \sum_{k=1}^{z_{ij}}(\beta S)^k\Big(c + \ln \mathcal{Q}_{0,j+k}^{-1}\Big) + \tau_{ij}. \tag{9}$$

Interpretation of the various terms in 9 is analogous to that for males. The main difference is that type j who marries a type i male agrees to the equilibrium transfer τ_{ij} from her partner.

Equilibrium Conditions and the measure of Total Marriage gains:

The formulation of the dynamic discrete choice problem outlined so far is standard. The approximation of the marriage market using this framework in the full commitment context is our own. It does not rely on the assumption that the divorce hazzard δ is constant across types and duration.**!!!!**[12] We diverge from the dynamic discrete choice literature in the way we take the restrictions represented by 9 and 6 to data. What follows outlines our empirical strategy.

Suppose we only observe aggregate number of matches by types and the aggregate number of available males and females by types. Let μ_{ij}^d be the number of type i males who want to marry type j females and μ_{i0}^d be the number of type i males who want to remain single. The maximum likelihood estimator for $\mathcal{P}_{ij}$ and $\mathcal{P}_{i0}$ are μ_{ij}^d/m_i and μ_{i0}^d/m_i respectively. The specification of the log-odds ratio from the male decision problem in equation (6) delivers a system of $(Z \times Z)$ quasi-demand equations for females of age j by age i males,

$$\ln\frac{\mu_{ij}^d}{\mu_{i0}^d} = \sum_{k=0}^{z_{ij}}(\beta S)^k\left(\frac{1}{2}\pi_{ij}^k - \alpha_{i+k,0}\right) - \sum_{k=1}^{z_{ij}}(\beta S)^k\Big(c + \ln\frac{m_{i+k}}{\mu_{i+k,0}^d}\Big) - \tau_{ij}, \quad \forall\ \ i,j = 1,\ldots,Z. \tag{10}$$

Similarly, if we use the maximum likelihood estimators for the probabilities in the female decision problem, we arrive at the following system of $(Z \times Z)$ quasi-supply equations,

$$\ln\frac{\mu_{ij}^s}{\mu_{0j}^s} = \sum_{k=0}^{z_{ij}}(\beta S)^k\left(\frac{1}{2}\pi_{ij}^k - \gamma_{0,j+k}\right) - \sum_{k=1}^{z_{ij}}(\beta S)^k\Big(c + \ln\frac{f_{j+k}}{\mu_{0,j+k}^s}\Big) + \tau_{ij}, \quad \forall\ \ i,j = 0,1,\ldots,Z. \tag{11}$$

At the marriage market equilibrium, the probabilities of matching for males and females and transfers τ_{ij} must be such that the number of i type men who want to marry j type spouses exactly equals the number of j type women who agree to marry type i men for all conbinations of (i,j). That is, for each of the $(Z \times Z)$ sub-markets,

$$\mu_{ij}^d = \mu_{ij}^s = \mu_{ij}$$

We apply the above marriage market clearing condition to the system of quasi-supply and demand equations. We add (10) and (11) to get the total gains to marriage for a couple marrying at age (i,j) when $z_{ij} > 0$ takes the form

$$\begin{aligned}\ln\frac{\mu_{ij}}{\sqrt{\mu_{i0}\mu_{0j}}} \quad &= \quad \frac{1}{2}\sum_{k=0}^{z_{ij}-1}(\beta S)^k\Big(\pi_{ij}^k - \alpha_{i+k,0} - \gamma_{0,j+k} - \beta S(2c + \ln\frac{m_{i+k+1}f_{j+k+1}}{\mu_{i+k+1,0}\mu_{0,j+k+1}})\Big) \\ &\quad + \frac{1}{2}(\beta S)^{z_{ij}}(\pi_{ij}^{z_{ij}} - \alpha_{i+z_{ij},0} - \gamma_{0,j+z_{ij}}).\end{aligned} \tag{12}$$

[12] The treatment in Section **??** generalizes this assumption.

Equation (12) says that the log ratio of the number of marriages between i males and j females to the geometric mean of the unmarrieds, is proportional to the systematic gains to i, j marriages relative to them remaining unmarried. We will use the total gains to an i, j marriage, $T_{ij} = \ln\big(\mu_{ij}/\sqrt{\mu_{i0}\mu_{0j}}\big)$. The first term of equation (12) is the discounted sum of net within marriage payoffs. Let $\alpha_{i+k,0}$ and $\gamma_{0,j+k}$ be the per period systematic gains for a male and female from being unmarried at age $i+k$ and $j+k$ respectively. Recall that dissolution of marriage occurs at an exogenous rate of δ. The parameter c is the Euler's constant which arised from our distributional assumption on the unobserved idiosyncratic deviations in marital output. In each period that a couple is locked into a match they forgo the utility from being single that period. This opportunity cost for an $i+k$ aged male is $\alpha_{i+k,0} + \beta S(c + \ln m_{i+k+1}/\mu_{i+k+1,0})$. It is comprised of a per period payoff of $\alpha_{i+k,0}$ plus the expected value of participating in the marriage market tomorrow captured by $\beta S(c + \ln m_{i+k+1}/\mu_{i+k+1,0})$. This latter term is proportional to the marriage rates of age $i+k+1$ males. If marriage rates of $i+k+1$ males is high, the expected discounted value of being single at age $i+k+1$ captured by $\beta S(c + \ln m_{i+k+1}/\mu_{i+k+1,0})$ would also be large. The opportunity cost for a female from being single at age $j+k$ is $\big(\gamma_{0,j+k} + \beta S(c + \ln(f_{j+k+1}/\mu_{0,j+k+1})\big)$. The interpretation of these terms are analogous to the male counterpart. Hence, the sum of these two terms gives the payoff to the couple from being single at ages $i+k$,$j+k$ and being able to participate in the marriage market.

So, for each period k of marriage, where $k < z_{ij}$, the systematic net gain from being married is

$$\pi_{i+k,j+k} - \alpha_{i+k,0} - \gamma_{0,j+k} - \beta S\big(2c + \ln \frac{m_{i+k+1} f_{j+k+1}}{\mu_{i+k+1,0}\mu_{0,j+k+1}}\big).$$

This gain is obtained only if the marriage remains intact which occurs with probability S^k. The last term of (12) is the systematic net gain in the last period of marriage z_{ij}.[13] Finally, when $z_{ij} = 0$ $T_{ij} = (\pi^0_{ij} - \alpha_{i0} - \gamma_{0j})$ because there are no futher marriage market conditions to consider. The bereaved spouse will re-enter the marriage market in the next period with certainty, and the next period is irrelevant for the age Z spouse.

If we compare the framework just outlined with the static model of Choo and Siow (2006), the authors showed that the total gains to marriage relative to remaining single in the static environment takes the form,

$$\ln \frac{\mu_{ij}}{\sqrt{\mu_{i0}\mu_{0j}}} = T_{ij} = \frac{1}{2}\big(\pi_{ij} - \pi_{i0} - \pi_{0j}\big) \tag{13}$$

Here, π_{ij} denote the total systematic discounted marital output from a marriage between an age i male and an age j female,[14] π_{i0} denote the systematic discounted gain to an age i male from not marrying at age i, and π_{0j} denote the systematic discounted gain to an age j female from not marrying at age j. (13) is intuitive. It says that in the static enviroment the log ratio of the number of marriages between i males and j females to the geometric mean of the unmarrieds, is proportional to the systematic gains to i, j marriages relative to them remaining unmarried. In that static model, single heterogenous males and females get one opportunity to match. One limitation of the static approach to the marriage market is that it ignores dynamic considerations. In particular, the value of delaying marriage at time

[13] In this last period, the older spouse does not have future opportunities to participate in the market, while for the younger spouse, he or she returns to the marriage market with certainty in the following period. In this case, the continuation values from participating in the market exactly cancel out.

[14] Systematic gains differ from expected gains by a constant.

t depends on future opportunities for marriage. Future opportunities are related to current population of single males and females $\boldsymbol{m}$ and $\boldsymbol{f}$ and the decisions that these individuals make. However if future opportunities affect the value of not marrying, then π_{ij} should be a function of these future opportunities and some exogenous preference parameters. Hence the form represented by equation 12 conforms to what we expect.

Generating Empirical Difference Equations:
Consider taking a quasi-difference of T_{ij} in equation (12). The assumption that δ is constant allows the quasi-difference to have the form,

$$T_{ij} \quad - \quad \beta S\ T_{i+1,j+1} = \frac{1}{2}\pi_{ij}^{0} - \alpha_{i0} - \gamma_{j0} - \beta S\Big(2c + \ln\frac{m_{i+1}f_{j+1}}{\mu_{i+1,0}\mu_{0,j+1}}\Big) + \frac{1}{2}\sum_{k=1}^{z_{ij}}(\beta S)^{k}\Big(\pi_{ij}^{k} - \pi_{i+1,j+1}^{k-1}\Big) \tag{14}$$

The special case where the time discount $\beta = 1$ and where individuals perceive their subjective divorce rate as zero or $S = 1$ is instructive. [15] In this case, for $z_{ij} > 0$, equation (14) becomes:

$$\ln\frac{\mu_{ij}}{\mu_{i+1,j+1}} - \ln\sqrt{\frac{\mu_{i0}\mu_{0j}}{m_{i+1}f_{j+1}}} = \frac{1}{2}\left(\sum_{k=0}^{z_{ij}}\pi_{ij}^{k} - \big(\alpha_{i0} + \gamma_{0j} + 2c + \sum_{k=1}^{z_{ij}}\pi_{i+1,j+1}^{k-1}\big)\right) \tag{15}$$

Without divorce, $\mu_{i0}\mu_{0j} \simeq m_{i+1}f_{j+1}$ and $\ln\sqrt{\frac{\mu_{i0}\mu_{0j}}{m_{i+1}f_{j+1}}} \simeq 0$. Deviations are only due to re-entry into the marriage market due to deaths of spouses. With positive assortative matching by spousal ages, this is not a concern for most i, j matches. So for $z_{ij} > 0$, $\beta = 1$ and $\delta = 0$ we have the following approximation of 14,

$$\ln\frac{\mu_{ij}}{\mu_{i+1.j+1}} = \frac{1}{2}\left(\sum_{k=0}^{z_{ij}}\pi_{ij}^{k} - [\alpha_{i0} + \gamma_{0j} + 2c + \sum_{k=1}^{z_{ij}}\pi_{i+1,j+1}^{k-1}]\right) \; = \; h(i,j) \tag{16}$$

The interpretation of (16) is as follows. $\sum_{k=0}^{z_{ij}}\pi_{ij}^{k}$ is the present value of the systematic return to an i, j couple marrying in the current period. $\alpha_{i0} + \gamma_{0j} + 2c + \sum_{k=1}^{z_{ij}}\pi_{i+1,j+1}^{k-1}$ is the present value of the systematic return to an i, j couple delaying their marriage for one year. Consider all marriages in a given period. $\ln\big(\frac{\mu_{i+1,j+1}}{\mu_{ij}}\big)$ is the growth rate of i, j marriages. It measures the systematic benefit to an i, j couple delaying marriage for one period rather than marrying immediately. This interpretation of the evolution of total gains in Equation (16) is similar to the evolution of pricing of capital goods in evolving competitive environments (e.g. Hansen and Sargent(2005)). Rather than pricing a durable commodity, we are pricing durable matches in evolving competitive environments.

Whether the above approximation is reasonable or not is an empirical issue which we will address in section 6. Assuming that there are Z different ages (or types) of males and females in the marriage market, there will be $2Z - 3$ of these first difference equations corresponding to spousal age gaps of $g = i - j$. From an empirical perspective, Equation (16) is useful because its left hand side is observable and its right hand side is only a function of exogenous parameters that are invariant to changes in supplies, $h(i, j)$. Thus, we can use parametric or semi-parametric techniques to estimate Equation (16).

[15] Surveys of young adults and newlyweds show that they perceive that their own divorce risks are far lower than that of the general population. Baker and Emery 1993; Mahar 2003; Weinstein 1980.

If $\ln\sqrt{\frac{\mu_{i0}\mu_{0j}}{m_{i+1}f_{j+1}}} \not\simeq 0$, then (15) applies. Now, it is easy to interpret the term $\ln\sqrt{\frac{\mu_{i0}\mu_{0j}}{m_{i+1}f_{j+1}}}$ in the left hand side of (15). Recall that the right hand side of (15) measures the systematic benefit to marrying now versus delaying for a period. If $\ln\sqrt{\frac{\mu_{i0}\mu_{0j}}{m_{i+1}f_{j+1}}} \not\simeq 0$, then the observed number of marriages $\mu_{i+1,j+1}$ not only reflects the number of age i males and age j females who delayed marriages for one period, but also includes marriages of $i+1$ males with $j+1$ females who married in the previous period and became single in the current period. These include the newly divorced as well as the newly bereaved. Thus, we need to deflate $\mu_{i+1,j+1}$ by $\sqrt{\frac{\mu_{i0}\mu_{0j}}{m_{i+1}f_{j+1}}}$ to get the correct number of single i,j couples who chose to delay marriage for a period. Empirically, it is as easy to estimate (15) as (16). The variation in the left hand side of (15) is dominated by the variation in $\ln\mu_{ij} - \ln\mu_{i+1.j+1}$.[16] So (16) is a reasonable empirical approximation of (15), and we will use it in this paper due to its empirical transparency.

The system of $2Z-3$ first order difference equations in (16) is not a complete specification of the behavioral model. We will need initial conditions and a system of accounting identities governing stock and flows of individuals in the system. The initial conditions arise when $z_{ij}=0$, that is,

$$T_{ij} = \ln\frac{\mu_{ij}}{\sqrt{\mu_{i0}\mu_{0j}}} = \frac{1}{2}\big(\pi^0_{ij} - \alpha_{i0} - \gamma_{0j}\big). \tag{17}$$

(17) gives terminal values for the difference equations (16), and there are $2Z-3$ of these terminal values.[17] In any stationary equilibrium, we also have the $2Z$ stock flow accounting identities,

$$\begin{aligned} m_1 &= \sum_{j=1}^{Z}\mu_{1j} + \mu_{10} \quad \text{when } i=1, \qquad m_1 = \mu_{i0} + \sum_{k=1}^{i}\sum_{j=1}^{Z-i+k} S^{i-k}\mu_{kj} \quad \forall\ i>1, \\ f_1 &= \sum_{i=1}^{Z}\mu_{i1} + \mu_{01} \quad \text{when } j=1, \qquad f_1 = \mu_{0j} + \sum_{k=1}^{j}\sum_{i=1}^{Z-j+k} S^{j-k}\mu_{ik} \quad \forall\ j>1. \end{aligned} \tag{18}$$

By repeated substitution of (15), we can express all marriages as a function of the unmarrieds μ_{i0} and μ_{0j}, and the exogenous parameters $h(i,j)$ and T_{ij},

$$\ln\mu_{ij} = \begin{cases} T_{ij} + \ln\sqrt{\mu_{i0}\mu_{0j}}, & \text{if} \quad z_{ij}=0 \\ \sum_{k=0}^{z_{ij}-1} h(i+k,j+k) + T_{i+z_{ij},j+z_{ij}} + \ln\sqrt{\mu_{i+z_{ij}0}\mu_{0j+z_{ij}}}, & \text{if} \quad z_{ij}>0. \end{cases} \tag{19}$$

Substituting in μ_{ij} from (19) and (17) into the accounting identities of (18), we will have a system of $2Z$ equations with $2Z$ unknowns, the unmarrieds μ_{i0} and μ_{0j}. These equations take T_{ij} and $h(i,j)$ as parameters. A stationary marriage market equilibrium consists of the solution to this system. [18] In general, we cannot solve for an equilibrium analytically. The empirical section of the paper provides numerical solutions based on different empirical specifications of the model. Our model rationalizes T_{ij}, and $h(i,j)$ is exogenous and invariant to changes in supplies. In the next section, we impose additional structure on T_{ij} and $h(i,j)$ to test various possible explanations of the marriage distribution. Equations

[16] In our data, the mean and standard deviation of $\ln\mu_{ij} - \ln\mu_{i+1,j+1}$ are 0.998 and 0.70 respectively. The mean and standard deviation of $\ln\sqrt{\frac{\mu_{i0}\mu_{0j}}{m_{i+1}f_{j+1}}}$ are -0.030 and 0.05 respectively. The correlation between $\{\ln\frac{\mu_{ij}}{\mu_{i+1,j+1}} - \ln\sqrt{\frac{\mu_{i0}\mu_{0j}}{m_{i+1}f_{j+1}}}\}$ and $\ln\frac{\mu_{ij}}{\mu_{i+1,j+1}}$ is 0.997. Thus treating $\ln\sqrt{\frac{\mu_{i0}\mu_{0j}}{m_{i+1}f_{j+1}}}$ as zero in $\{\ln\frac{\mu_{ij}}{\mu_{i+1,j+1}} - \ln\sqrt{\frac{\mu_{i0}\mu_{0j}}{m_{i+1}f_{j+1}}}\}$ is reasonable.

[17] There are two terminal values for z_{Z1} and z_{1Z} which are not associated with any difference equations.

[18] The reduction of the number of endogenous variables to $2Z$ unmarrieds is similar to CS.

(19) and (17), together with the accounting identities of (18), define a semi parametric model of the marriage distribution without imposing structure on per period marital output.

It is convenient here to take stock of the empirical content of our model. The model consists of three components: A system of growth rate equations, equations (16). A set of terminal values for those growth rate equations, equations (17). A system of accounting identities, equations (18). The accounting identities have to hold independent of any theory of the marriage market. They are not specific to our particular model. Our model implies that the marriage distribution can be modeled by its growth rates as functions of exogenous parameters without restricting the behavior of those growth rates. However, the growth rates as specified by (16) are not a complete empirical model of the marriage distribution because it does not have terminal values. Our theory implies that (17) is an appropriate way to construct terminal values. But other static marriage functions discussed in the literature can also be used to construct terminal values.

In this paper, a distinct marital match is characterized by the spousal age gap, $i-j$. The generalization to include other fixed spousal characteristics such as ethnicity, religion, etc., is immediate. While a semi-parametric model of the growth rates of marriages can fit the marriage distribution, it cannot explain the empirical regularities in marriage distributions discussed in the introduction of this paper. We will explore restrictions on marital output in the next section to rationalize these regularities.

3 Identifying Preferences : Separable spousal contributions and marriage specific capital

What current gains to marriage, π_{ij}^k, explain the observed structure of T_{ij} and μ_{ij}? We explore a simple decomposition of π_{ij}^k into separable spousal contributions and marriage specific capital:

$$\pi_{ij}^k = \alpha_{i+k} + \gamma_{j+k} + \kappa_k + \nu_{ijk} \tag{20}$$

α_{i+k} and γ_{j+k} are the husband's and wife's contributions respectively. Each spousal contribution depends only on the spouses own age at the time of the contribution. There is no spousal age interaction effect. So, we explicitly rule out preferences for own aged spouses, the usual modeling construction used to generate positive assortative matching in spousal ages. κ_k represents the value of current marital output that is generated by the stock of marriage specific capital. This stock, κ_k, depends on how long the marriage has lasted. ν_{ijk} is a mean zero (estimation) error which is orthogonal to all other preference parameters, that is $\mathbb{E}(\nu_{ijk} \mid \forall\ i, j, k) = 0$.

Using (16) and (20), we get

$$\ln \frac{\mu_{ij}}{\mu_{i+1,j+1}} = \varphi + D_i\widetilde{\alpha}_i + D_j\widetilde{\gamma}_j + D_{z_{ij}}\widetilde{\rho}_{z_{ij}} + \psi_{ij}, \tag{21}$$

where D_i and D_j are dummy variables that take a value 1 when the male and female ages are i and j respectively and zero otherwise. D_t is a dummy variable that equals 1 at the maximal length of marriage, that is $t = z_{ij}$, and ψ_{ij} is a composite error that is uncorrelated with the covariates in (21). The terms $\widetilde{\alpha}_i$ and $\widetilde{\gamma}_j$ estimate the male and female relative contributions, $\frac{1}{2}(\alpha_i - \alpha_{i0})$ and $\frac{1}{2}(\gamma_j - \gamma_{0j})$ respectively.

We cannot expect to identify all the separate male and female relative contributions to marriage. The reason is that every marriage has a husband and a wife, a fixed coefficient production function. Thus, we identify the separate relative contributions to marriage by each gender up to the normalizations,

$$\sum_{i=1}^{Z}(\alpha_i - \alpha_{i0}) = 0, \quad \text{and} \quad \sum_{j=1}^{Z}(\gamma_j - \gamma_{0j}) = 0.$$

$\widetilde{\rho}_t$ is proportional to the the expected discounted value of increments in marriage specific capital for a marriage of maximum length $t = z_{ij}$. Substituting the spousal contribution equation (20) into the growth rates equation, we get $\widetilde{\rho}_t = \frac{1}{2}\sum_{k=0}^{z_{ij}}(\kappa_k - \kappa_{k-1}) = \frac{1}{2}\kappa_{z_{ij}}$ assuming that $\kappa_{-1} = 0$. Thus the coefficient on the dummy variable D_t is proportional to the stock of marriage specific capital for a marriage of length z_{ij}. The identification of κ_l by $\widetilde{\rho}_t$ may be motivated as follows. If an i, j marriage in the current period can last at most z_{ij} periods, then an $i+1, j+1$ marriage next period can last at most $z_{ij} - 1$ periods. So the marriage next period also gets the returns to marriage specific capital for all periods up to $z_{ij} - 1$. Thus the gain in marriage specific capital from marrying this period instead of delaying one period is the return to the stock of marriage specific capital for period z_{ij}, that is $\kappa_{z_{ij}}$. φ is equal to c plus the sum of the average male and female relative contributions to marriage. Although the parameters of regression equation (21) have a structural interpretation, there is little apriori structure imposed on the data. One may interpret (21) as a semi-parametric description of the evolution of the marriage distribution.

3.1 Terminal values and other identification issues

In order to simulate the model, we need to estimate terminal values for the difference equations (21). The separable spousal contribution model implies that at $z_{ij} = 0$, the terminal value given by (17) has the form

$$\ln\frac{\mu_{ij}}{\sqrt{\mu_{i0}\mu_{0j}}} = \frac{1}{2}(\widetilde{\alpha}_i + \widetilde{\gamma}_j + \kappa_0 + v_{ij0}) \tag{22}$$

The first differenced regression (21) identifies $\widetilde{\alpha}_Z$, $\widetilde{\gamma}_Z$ from estimates of $\widetilde{\alpha}_{Z-1}$ and $\widetilde{\gamma}_{Z-1}$. Using (22), let $\widetilde{T}_{ij} = \frac{1}{2}(\widetilde{\alpha}_i + \widetilde{\gamma}_j + \kappa)$ for $z_{ij} = 0$ where κ is a free parameter. Given κ, the model can be used to simulate a predicted marriage distribution. We choose κ to match the total number of predicted marriages to the actual number of total marriages.

Another method to obtain estimates of terminal values is to use the observed data on marriages and unmarrieds, $\ln(\widehat{\mu}_{ij}/\sqrt{\widehat{\mu}_{i0}\widehat{\mu}_{0j}}) = \widehat{T}_{ij}$, as an estimate for T_{ij} for $z_{ij} = 0$. $\widehat{T}_{ij}$ does not impose the separable spousal contribution model at the terminal values. The main drawback of using $\widehat{T}_{ij}$ to estimate T_{ij} is that there are few marriages at the last age for males or females. Therefore the estimates of T_{ij} are imprecise. Since we use these estimates as initial values for our simulations, whatever imprecision they contain is propogated through the simulated model. Even in this case, we will need to add a free intercept κ to $\widehat{T}_{ij}$ to match the total number of predicted marriages to the actual number of total marriages.

Importantly, note that our parametrization of marriage specific capital is not observationally equivalent to a per period preference for own type matches. Consider an alternative parametrization of π^k_{ij}:

$$\pi^k_{ij} = \alpha_{i+k} + \gamma_{j+k} - H(\ln i - \ln j) \tag{23}$$

where $H(.)$ is an increasing continuous function of $\ln i - \ln j$. Notice that the contribution of own type gain to π_{ij}^k depends on the distance between i and j. Holding the age of the older spouse constant, varying the age of the younger spouse will change the total contribution of own type gain to π_{ij}^k. But holding the age of the older spouse constant, varying the age of the younger spouse will not change the total contribution of marriage specific capital to π_{ij}^k because z_{ij} is unchanged. As such, marriage specific capital is not observationally equivalent to a per period gain for own type matches.

Another possible concern is if the accumulation of marriage specific capital is important for marital matching and women live longer than men, should we not observe women marrying younger men? Due to differential mortality, women are on the long side of the marriage market. Since the death hazard is increasing in age, individuals will prefer younger spouses especially when the accumulation of marriage specific capital is important. Since men are relatively scarce in the marriage market, men can marry younger wives without having to pay a large transfer. Thus the observation that men marry younger women is consistent with marriage specific capital being an important consideration in the choice of a spouse.

We deal with differential mortality in two ways. First, we simulate versions of the model adjusting the accounting identities to reflect the actual mortality schedules for men and women. Second, differential mortality may also show up as a preference for a younger spouse. In order to allow for potential differential mortality effects, we also consider empirical models where we add $H(.)$ to (21). As discussed above, $H(.)$ is in general identified even in the presence of marriage specific capital.

Recall that $\widetilde{\rho}_t$ in (21) is proportional to the expected discounted increment in marriage specific capital for marriage of maximum length $t = z_{ij}$ where $z_{ij} = \min(Z - i, Z - j)$. When life expectancy depends on current age and gender, we may approximate $z_{ij} = \min(\ Z_i^m - i, Z_j^f - j)$ where Z_i^m is life expectancy of an age i male and Z_j^f is the life expectancy of an age j female. Assume that $Z_j^f \geq Z_i^m$, then

$$\begin{aligned} z_{ij} &= \min\{Z_i^m - i, Z_j^f - j\} \\ &= Z_i^m - \max\{i, j - (Z_j^f - Z_i^m)\} \\ &= Z_i^m - \max\{i, j\} - \Phi(j - i, Z_j^f - Z_i^m), \end{aligned}$$

where

$$\Phi(j-i,k) = \begin{cases} 0 & \text{if} \quad i \geq j, \\ j - i & \text{if} \quad j - k \leq i < j, \\ k & \text{if} \quad i < j - k. \end{cases}$$

Thus, we approximate z_{ij} as:

$$z_{ij} = z_i + z_j - \max(i, j) - H(\ln i - \ln j)$$

So adding $H(.)$ to (21) allows for gender specific mortality and marriage specific capital. In this more general model. $\widetilde{\alpha}_i$ and $\widetilde{\gamma}_j$ can no longer be associated solely with spousal contributions. We also cannot distinguish between a concern for own type per se versus marriage specific capital and gender specific mortality. Without loss of continuity, readers who are primarily interested in empirical work may skip to section (4).

4 Data

To construct the marriage distribution by age, $\widehat{\mu}_{ij}$ for 1990, we use the Vital Statistics [19] files from the NBER collection of the National Center for Health Statistics. These files contain a sample of the marriage records from reporting states. We use the weights in the Vital Statistics files to get an estimate of the total number of each type of marriages in reporting states.[20] $\widehat{\mu}_{ij}$ is equal to our estimate of the number of marriages in reporting states in 1991 and 1992 divided by two.

The number of unmarrieds by age, $\widehat{\mu}_{i0}$ and $\widehat{\mu}_{0j}$, come from the 5% 1990 US census. We included all unmarried individuals in reporting states, older than 15 and younger than 77. Unmarried means marital status not equal to (i) married spouse present, or (ii) married spouse absent. We used the Census weights to get an estimate of the total unmarried counts. Summary statistics of the data are given in Table 1 below.

Table 1

A: Summary of marriages 90/91, Vital Statistics

	Males	Females
Mean Age	31.1	28.8
Std Deviation	10.4	9.4
Median Age	27.8	25.8
Average age difference	2.39	
Average # of marriages	1817844	

B: Summary of unmarrieds, 1990 US Census

	Males	Females
Total # of unmarrieds (mil.)	31.36	35.38
# of 16 year olds (mil.)	1.433	1.368
Mean Age	31.9	38.4
Std Deviation	14.7	18.6
Median Age	26.7	32.6

Table 1A shows summary statistics of the marriage data that we use. The mean age of new marriages from the reporting states in 1990/91 for males and females are around 31 and 29 years respectively.[21] While the median age for males and females is around 26 and 28 years respectively. Like the mean,

[19] This data is collected by the US Department of Health and Human Services.

[20] Alabama, Alaska, California, Colorado, Connecticut, Delaware, District of Columbia, Florida, Georgia, Hawai, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, New Hampshire, New Jersey, New York State, North Carolina, Ohio, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, Utah, Vermont, West Virginia, Wisconsin, Wyoming,

[21] Since most single adults marry relatively young, the marginal distribution by gender is skewed to the right generating median ages that is lower than the mean.

the gap in the median age of marriage between males and females is around two years. The average number of reported new marriages for 90/91 is around 1.8 million and the average age gap between these newly-wed couples is around 2.4 years.

Table 1B provides some summary statistics of the 5 % 1990 US census. The number of single males and females in the reporting states in 1990 are around 31.4 and 35.4 million respectively. While the number of 16 year old single males is around 1.43 million compared to 1.37 million single 16 year old females. That is, there are around 4.7 % more males than females at 16. The mean age of single females is around 38 compared to 32 for males. This difference is largely due to the increase in the single old female after the age of 50 as shown in Figure 1b. The standard deviation for females is also larger than that for males. Additional statistics of the marriage distribution are provided in the first column of Table 3 and 4. These statistics we try to match using our model.

Figure 1c shows the marriage rates for men and women by age. As expected, except for the youngest ages, the marriage rates decline by age. Finally, Figure 1d shows the estimated total gains to marriage, equation (13). As documented earlier in CS, the total gains distribution shows that the total gains to marriage are highest for young adults who are married to each other. The decline in total gains as adults age is more gradual than that shown in Figure 1a.

5 Empirical results

Estimators of our dependent variable, $T^t(i,j)$ are constructed using data on marriages and supplies of unmarrieds. For marriages where the age of spouses, i and j are far apart, μ_{ij}^t is typically a very small number making the variance of estimator $\widehat{T}^t(i,j)$ very large relative to observations where i and j are close. This thinness of cell problem (partly alleviated though aggregation across ages and time) makes accounting for sampling error for these estimators very important. Ignoring these variances would mean that the least squares gives equal weights for all observations irrespective of the distance between i and j. We provide two set of standard errors. The first is the robust White's standard error implied by the orthogonality condition in least squares, and the second is the efficient standard error according to the limiting variance-covariance matrix derived from the multinomial assumptions of the model. This limiting distribution is derived in Appendix A.

Our model depends on three components: (1) A model of the growth rates of marital matches. (2) A related model for terminal gains. And (3), a system of accounting equations. In this section, we estimate component (1), models of the growth rates of marital matches.

Figure 2 shows the smoothed bivariate distribution of marriage growth rates. Figure 2a shows the fitted growth rates according to our preferred model (more on this later) while Figure 2b shows the smoothed estimates from the raw data. Most of our data lie along the age diagonal. The off diagonal surface in Figure 2b is due to smoothing of sparce data. The variance of our non-parametric estimates increase when looking at growth rates of marriages between old men and young women and vice-versa. The fitted estimates from model 2 appears to approximate the growth rates between closely aged couples reasonably well. The structure in model 2 uses this information to approximate the growth rates where the data is scarce. The distribution of growth rates is quite different from the marriage distribution in Figure 1. Ignoring observations with zero marriages, we have altogether 2512 growth rate observations. The mean growth rate is 0.007 with a standard deviation of 0.716. We estimate the three growth rate models (21) by ordinary least squares. Table 2 present summary statistics for these estimated models. Appendix A contains the point estimates.

Table 2: Marriage growth rate regressions

	Model		
	1	2	3
Gender specific spousal age effects	√	√	√
Marriage specific capital, z_{ij}, effects		√	√
Concern for spousal age difference, $H(\ln i - \ln j)$ polynomial			√
p value for gender neutral spousal age effects		0.062	0.058
p value for z_{ij}		0.000	0.000
p value for $H(\ln i - \ln j)$			0.994
Number of observations	2512	2512	2512
R^2	0.1489	0.1742	0.1743

Column (1) is the benchmark model. It has dummy variables for husband's and wife's ages at marriage but no dummy variable for marriage specific capital, z_{ij}. Column (2) has dummy variables for husband's and wife's ages at marriage and dummy variables for marriage specific capital. Since we calculate robust standard errors, we test all restrictions using asymptotic Wald tests. The p-value for testing whether all the dummy variables for marriage specific capital are zero is 0.000. Thus we easily reject the hypothesis that there is a concern for marriage specific capital. The p-value for testing whether a husband's contribution at age i is equal to a wife's contribution of the same age (i.e. gender neutral spousal age effects) for all ages, is 0.062. Gender differences in spousal contributions may be of a second order in explaining marital growth rates. Column (3) adds three polynomials in spousal age differences, $(\ln i - \ln j)^2$, $(\ln i - \ln j)^3$ and $(\ln i - \ln j)^4$ to capture the concern for spousal age differences. As discussed earlier, we may interpret the presence of spousal age differences as a concern for marriage specific capital when there is gender specific mortality. The p-value that the coefficients on these three variables are jointly equal to zero is 0.994. Thus, there is little evidence that, after controlling for marriage specific capital, adding a concern for spousal age difference matters in explaining marital growth rates. The p-value for testing whether all the dummy variables for marriage specific capital are zero remains at 0.000. The p-value for testing whether a husband's contribution at age i is equal to a wife's contribution at the same age for all ages, is 0.058. Thus, the results in Table 1 show that the separable spousal contribution model with marriage specific capital (column (2)) is a reasonable description of marital growth patterns. A note of caution is necessary. The R^2 of the regression is 0.174. So there is substantial variation in the growth patterns that we are not capturing. How important this residual variation is will be shown in the simulations.

We now discuss the point estimates of column (2). Because there are so many point estimates, Figure 3a presents a plot of the point estimates of the spousal contributions. Recall that the means of each spousal contribution is normalized to zero. Figure 3b shows the smoothed spousal contributions estimates while Figures 3c and 3d show these estimates with their respective smoothed 95% confidence intervals. These plots show that the precision of these spousal contribution estimates worsen with age, with the estimates below the age of 30 being the most precise. The spousal contributions rise steeply with age and stablize after the age of 30. It again increases at the end of the lifecycle, but the standard errors on these estimates are large. The female spousal contributions are higher than male spousal contributions at the early ages and the pattern reverses in the early twenties. This gender asymmetry is potentially important in generating the younger average age at marriage for women. It is also implied by fertility models of marriage.[22] By age 50, male spousal contributions are marginally higher than female spousal contributions but the differences are neither economically nor statistically significant.

That the estimated spousal contributions are low for young adults is surprising because marriage rates for young adults are generally larger than for older adults. We will provide an explanation for this surprising finding next.

Figure 4a shows the point estimates on the dummy variables D_t, the potential maximum lengths of marriage. Figure 4b plots the smoothed estimates together with the smoothed 95% confidence intervals. Interpreting the estimates as the stock of marriage specific capital in period z_{ij}, $\widetilde{\rho}_t = \frac{1}{2}\kappa_{z_{ij}}$, the figure shows that the stock of marriage specific capital, κ_t, increases early with t, stabilizes, and then increases again at the end. This steep increase after $z_{ij} > 45$ is needed to fit the strong positive assortative matching by age among young adults. A young individual who marries an older spouse will anticipate having less time to accumulate and enjoy marriage specific capital relative to marrying a same age or younger spouse. From this perspective, younger spouses prefer to marry each other. The steep increase is also responsible for the high marriage rates among young adults. Individuals can only capture the steep increase by marrying young. This steep increase mitigates the need for spousal contributions to be higher for young adults to capture their higher marriage rates. When the stock of marriage specific capital increases strongly with age, there is an incentive to marry early. The estimates for marriage capital stock at the early years of marriage (when z_{ij} is small) are more imprecise than at later years of marriage (when z_{ij} is large). The identification of marriage capital stock at the later years of marriage relies on the growth rates of marriages of young couples where most of our data lies. The marriage capital stock at the early years is identified off new marriages between old couple or couples where one of spouse is much older than the other. These observations in the data are relatively scarce accounting for the larger standard error on marriage capital estimates at the early years of marriage. While the smoothed confidence interval for each point estimate is large, recall that the p-value for $z_{ij} = 0$ for all i, j was strongly rejected.

A point of caution is needed. We interpreted the estimated parameters in Figures 3 and 4 as parameters which determine per period spousal output. Economists do not have much experience estimating per period spousal output. Thus, whether our parameter estimates are reasonable from a behavioral perspective is speculative.

5.1 Stationarity assumption and accounting identities

Our model depends on three components: (1) A model of the growth rates of marital matches. (2) A related model for terminal gains. And (3), a system of accounting equations. We have estimated parameters for components (1) in the previous section. As discussed in section 3.1, we can obtain estimates of terminal gains given the estimates of the growth rates model.

How well our model approximates the observed marriage distribution also depends on our specification of the accounting equations. Our accounting equations may be misspecified due to non-stationarity,

[22] E.g. Hamilton and Siow; Siow 1998; Giolito.

preference shocks, duration dependent divorce hazards, immigration and emmigration in the society and other factors.

In order to get a quantitative assessment of how well our model can potentially fit, we predict the number of unmarried men and women by age using the observed marriage distribution, the gender specific mortality rates and gender specific entry cohort sizes, and our accounting equations. This provides a benchmark in the best case scenario when we model the growth rates of marriage perfectly. The divorce rate for the accounting equations, 1.63% per annum, is estimated by minimizing the sum of squared errors between the simulated unmarried and observed unmarried distributions.

Figure 5 shows the simulated and observed unmarried distributions. The discrepancies between the simulated and actual unmarried distributions are due to misspecification of the accounting equations. Thus, Figure 5 shows in the best case scenario, how well a stationary model such as ours can do in rationalizing the observed distribution of unmarrieds.

Three points are important to stress. First, the qualitative difference between simulated and observed male and female unmarried behavior is the same. There is a relative surplus of unmarried young males relative to unmarried young females and the reverse is true for older individuals. Second, there are more observed unmarried middle aged adults and fewer observed unmarried older adults than in the simulations. Lastly, we observe that the observed unmarrieds at the last age is significantly lower than the respective simulated unmarrieds. All the simulations in this paper require an initial condition given by the marriage gains at the terminal age of marriage. However, using the empirical terminal gains (with the observed unmarrieds) as the initial condition, would be an over-estimate the terminal gains implied by the best fitting model. Hence, the intercept κ in terminal gains is used to scale this discrepancy in observed and theoretical initial conditions. In our simulations, the free parameter κ is used to match the total number of simulated marriages with the observed number of marriages.

Table 3 : Descriptive Statistics from various models

	Data	Models				
		1	2	3	1A	2A
Gender specific spousal contributions		√	√	√	√	√
Marriage specific capital			√	√		√
Concern for spousal age difference				√		
Observed terminal gains						√
Gender specific mortality		√	√	√	√	√
Gender specific entry sizes		√	√	√	√	√
Male mean age of marriage	31.14	37.44	33.98	34.55	42.28	33.42
Female mean age of marriage	28.76	37.62	31.76	31.86	37.61	30.47
Male SD age of marriage	10.3	13.16	11.07	11.14	11.11	10.13
Female SD age of marriage	9.4	12.59	10.15	10.14	10.41	9.14
Correlation of spousal ages	0.83	0.31	0.7	0.62	0.85	0.89
Average spousal age difference	2.39	-0.18	2.22	2.68	4.67	2.95

6 Rationalizing the marriage distribution

Table 3 presents simulations of the marriage distribution based on different models. These models vary the structures put on spousal contributions in the current period return to marriage. This allows us to test the importance of different theories of marital outputs in rationalizing the marriage distribution.

The first column presents summary statistics for the observed marriage distribution. The mean ages of marriage for men and women are 31.14 and 28.76 respectively. The average spousal age difference (husband's age minus wife's age) is 2.39 years. The standard deviations of marriage ages for men and women are 10.35 and 9.37 respectively. The correlation in spousal ages at marriage was 0.83.

Columns 1, 2 and 3 present simulation results corresponding to Models 1, 2 and 3 discussed in Table 1. We begin with a discussion of the simulations from our preferred model, Model 2 (column 2). The model has gender specific spousal contributions, accumulation of marriage specific capital in the estimated first differenced equation, and estimates of terminal gains constructed from estimated spousal contributions, T^1_{ij}. It allows for gender specific mortality and gender specific cohort entry sizes in the accounting equations.

The mean ages for marriage are 33.98 and 31.76 for men and women respectively. They are 9.1% and 10.0% higher than the respective observed values. The standard deviations of marriage ages are 7.5% and 8.0% higher than the observed values for men and women respectively. As shown in Figures 6a and 6b, the simulated marginal distributions of marriage ages are shifted to the right and have larger dispersion than the actual distributions. Figures 7a and 7b show the simulated unmarrieds distributions. We underpredict the number of unmarrieds younger than 50 and overpredict those older than 50. These discrepancies are similar qualitatively, but smaller quantitatively, than those generated by the unmarried distributions in Figure 5 using the actual marriage distribution. Thus we believe that the discrepancies in fit of both the simulated marriage and unmarried distributions are primarily due to our many stationarity assumptions. In terms of marriage matching patterns, the simulated correlation in spousal marriage ages is 0.70, which is 16% lower than the observed correlation. The simulated average spousal age gap is 2.22 years, which is 7.1% lower than what is observed. Figure 8 compares the observed marriage distribution with that predicted according to model 2. It is clear that our model of separable spousal contributions with tenure capital provides a reasonable approximation of the marriage distribution. The ripples along the diagonal in the plot of Figure 8b also highlight the working of the growth rates of marriage used by our model.

Other than the divorce rate and the intercept in terminal gains, κ, all other parameters of our model are estimated from a single marital growth rate regression model. A growth rate equation does not restrict the level of marital matching at a point in time. To generate cross sectional marriage matching patterns, we need to obtain the total number (level) of marriages for each growth rate equation at at least one point in time. Total gains from marriage, including terminal gains, contain information on the levels of marital matching. Our estimates of terminal gains, $\widetilde{T}_{ij}$, are constructed with only parameter estimates from the growth rate regressions. This construction of $\widetilde{T}_{ij}$ is only feasible based on a structural interpretation of our model. The simulated marriage matching patterns are entirely generated by parameter estimates from the growth rate regession, κ, and the accounting equations.[23]

[23]In general, if we want to simulate a system of N first order difference equations, we need N initial conditions. The simulated model in column 3 estimates one initial condition parameter, κ. The rest are obtained from the estimates of

Model 3 (column 3) differs from model 2 by adding a concern for spousal age in the estimated first differenced equation. The mean ages for marriage are 34.55 and 31.86 for men and women respectively. They are 9.5% and 10.8% higher than the respective observed values. The standard deviations of the ages of marriage are 8.2% and 7.9% higher than the observed values for men and women respectively. In terms of marriage matching patterns, the simulated correlation in spousal marriage ages is 0.62, which is 25.0% lower than the observed correlation. The simulated average spousal age gap is 2.68 years, which is 12.1% higher than that observed. Although model 3 is more general than model 2, the fit of the simulated marriage distribution is worse in all dimensions except for the average spousal age difference. A reason for this poorer performance is that our more general model may be overfitting the growth rate data. Recall that the p-value for the extra parameters in the growth rate regression that were not needed was 0.994. Thus, both, the growth rate regression and simulation results suggest that, after controlling for a concern for marriage specific capital, adding a concern for spousal age difference is empirically redundant.

From hereon, we ignore the concern for spousal age difference. How important is the concern for marriage specific capital? Model 1 (column 1) is the same as model 2 except that it does not allow for marriage specific capital. As shown, the fit of the simulated marriage distribution is substantially worse than that in Model 2. In particular, the correlation in spousal ages fall to 0.31 and the average spousal age difference is -0.18. In otherwords, a separable spousal contribution model of marital output alone cannot replicate the distinctive features of the marriage distribution. This worse fit also accords with the theory. Without a concern for marriage specific capital, there should less positive assortative matching in spousal ages.[24]

The simulation results of model 1 are not a comparative statics exercise. To do a comparative statics exercise, we simulated the model with estimated spousal contribution parameters from model 2 but with all mariage specific capital parameters set to zero.[25] The results are shown in the column labelled Model 1A. The average ages of marriage for both men and women are substantially larger than the observed averages. Put another way, marriage specific capital is quantitatively important for explaining why most adults marry when they are young.

To explore the advantage that information on marriage patterns can bring, we simulate model 2A which differs from model 2 by using $\widehat{T}_{ij}$ at $z_{ij} = 0$ as the estimates of terminal gains rather than $\widetilde{T}_{ij}$. As discussed above, $\widehat{T}_{ij}$ retains the general lifecycle model as defined by equation (12) but it does not impose the separable spousal contribution model, equation (20).

The mean marriage ages are 33.42 and 30.47 for men and women respectively. They are 7.0% and 5.9% higher than the respective observed values. The standard deviations of marriage ages are 1.7% and 2.3% lower than the observed values for men and women respectively. In terms of marriage matching patterns, the simulated correlation in spousal ages of marriage is 0.89 which is 7.2% higher than the observed correlation. The simulated average spousal age gap is 2.95 years which is 23.0% higher than what is observed. Except for the average spousal age gap, model 2A matches the observed marriage distribution better than model 2. Thus, model 2A shows that there is a significant quantitative advantage to bringing information on marriage matching patterns, as summarized by $\widehat{T}_{ij}$, to the simulations.

our growth rate model.

[24]No preference for positive assortative matching does not imply a zero correlation in spousal ages because the numbers of available men and women by age are highly correlated.

[25]κ is re-estimated to match total observed marriages.

Table 4 : Supply Effects

	Data	Models			
		4	5	6	7
		10	7	8	9
Gender specific spousal contributions			√	√	√
Marriage specific capital		√	√	√	√
Concern for spousal age difference					
Observed terminal gains					
Gender specific mortality		√		√	
Gender specific entry sizes		√			√
Male mean age of marriage	31.14	34.14	34.28	34.04	34.23
Female mean age of marriage	28.76	33.76	32.14	31.66	32.21
Male SD age of marriage	10.3	11.04	11.39	11.1	11.36
Female SD age of marriage	9.4	10.99	10.38	10.07	10.42
Correlation of spousal ages	0.83	0.59	0.66	0.71	0.66
Average spousal age difference	2.39	0.384	2.14	2.39	2.02

We should not be surprized by this advantage. As discussed earlier, the model 2A may be viewed as a semi-parametric growth rate model of the marriage distribution without imposing any behavioral interpretation on the parameter estimates.

6.1 Relative scarcity and gender differences in mortality

The birth rate in the US for males is slightly higher on average than for females.[26] However, mortality hazard increases much faster with age for men. Even with more boys than girls being born, women still outnumber men in the United States. In 2003, the Census Bureau estimated a total of 144,513,361 females of all ages, compared to 138,396,524 males. As women have longer life expectancies than men, older men are scarce relative to older women in the marriage market. This relative scarcity of older men might advantage them in the marriage market providing them with better marriage possibilities with younger women. We want to ask the question, what would the marriage distribution have been like if there were no gender differences in both mortality and birth rates. What would be the extent of assortative matching without these gender differences? Models 4 to 7 in Table 4 explore the importance of these supply effects.

According to the 1990 US Census used in our empirical application, the population of 16 year old males and females was estimated to be around 1.433 and 1.368 million individuals respectively (as shown in Table 1B). That is, there are around 4.7% more 16 year old single males than females. We also use gender mortality rates from the US 89-91 Decennial Life Tables. These tables give the proportions

[26]For example, according to the National Vital Statistics Report, 2005, the highest sex birth ratio occurred in 1946 (1,059 male births per 1,000 females) while the lowest occurred in 1991 and again in 2001 (1,046 male births per 1,000 females).

of persons alive at the beginning of an age interval dying during that interval. For example, among the males and females who celebrated their 29th birthday, 0.19% of males and 0.07% of females would not survive to their 30th birthday, that is the mortality rate for males is 2.7 times larger than than of females. This gap actually narrows as these rates increase exponentially with age. For individuals who survived to their 49th birthday, the mortality rate for males increases to 0.58% and 0.32% for females, and by their 60th birthday, the rate for males is 1.46% and for females is 0.82%.

In model 4, we consider a symmetric spousal contribution model with marriage specific capital. That is, we estimated a gender neutral spousal contribution model with marriage specific capital. So, any gender differences in the marriage distribution are due to gender differences in supplies. In this symmetric enviroment, supply effects causes the average age of marriage for males to be around 0.4 years higher than for females. The effect of the higher male mortality rate dominates, making older men relatively more scarce than younger men. This relative scarcity advantages older men by lowering the transfers required from older men relative to younger men. This provides better marriage possibilities with younger women generating the differential in average age of marriage. The average spousal age difference falls to 0.384 years. In otherwords, with marriage specific capital, two years of the observed average spousal age gap are due to gender differences in the age profiles of spousal contributions.

Models 5, 6 and 7 employs the same marital output model as Model 2. Model 5 shuts down gender specific mortality and entry cohort sizes. The simulated model fits worse than Model 2 in all dimensions. The correlation in spousal ages is 0.66, and the average spousal age difference is 2.14. Comparing the fit with model 1, supply effects of the quantitative magnitudes considered here are less important than marriage specific capital in rationalizing marriage matching patterns.

Compared with the model 2, model 6 shuts down gender specific entry cohort sizes, and model 7 shuts down gender specific mortality. Comparing model 5 with model 6, adding gender specific mortality reduces the relative supply of older men. As expected, this increases the average spousal age difference. Every statistic in model 6 is quantitatively closer to the observed value than the equivalent statistic in model 5. That is, adding gender specific mortality to the model leads to a uniform improvement to the fit of the model.

Comparing model 5 with model 7, adding gender specific entry cohort sizes increases the relative supply of younger men (the opposite of model 6), reducing the average spousal age difference. Thus, the two supply effects have opposite effects on the average spousal age difference.

We have also explored models corresponding to columns 4 to 7 using $\widetilde{T}_{ij}$ as our estimate of terminal gains rather than $\widehat{T}^1_{ij}$. These models provide a slight quantitative improvement in terms of fit over the spousal contribution model similar to that between rows 2A and 2 in Table 3. Within the $\widetilde{T}_{ij}$ class, the ranking of the different features in fitting the marriage distribution is similar to that in columns 4 to 7.

Summarizing our empirical investigation, gender specific spousal contribtions and marriage specific capital are important determinants of the marriage distribution. Concern for spousal age is quantitatively unimportant. While both supply factors are significant, gender specific mortalities are quantitatively more important than gender differences in initial supplies.

In terms of overall fit, our simple stationary model can replicate well know features of the marriage distribution. We overpredict the average ages of marriage. Since we solve for the marriage distribution from the end, from an accounting point of view, we underpredicted the rate of decline in marriages for older adults. That is, we overpredicted per period marital output for older adults. We also underpredict the number of middle aged unmarrieds and over predict the number of older unmarrieds. These last discrepancies are due to our stationarity assumptions.

7 Conclusion

We propose and estimate a lifecycle model of the marriage distribution. The estimated model is able to rationalize the distinctive features of the marriage distribution.

There is much room for futher work. The estimated model misfits the marriage distribution in some dimensions. It will be important to improve the model in these dimensions. We have estimated a stationary model and ignored problems of time varying preferences, demographic shocks and duration dependent divorce hazards. These issues need to be addressed.[27] Adding matching in other attributes other than spousal age is important. In particular, adding an educational decision will be a significant advance. We need to incorporate endogenous divorce. The challenge is to add endogenous divorce and retain the empirical simplicity of the current framework. We have also ignored the issue of learning about the marital match and unobserved heterogeneity. These are significantly more difficult issues.

References

[1] Aguirregabiria, Victor and Pedro Mira."Sequential Estimation of Dynamic Discrete Games," *Econometrica*, forthcoming.

[2] Anderberg, Dan. "Getting Hitched: The Equilibrium Marriage Market Behaviour of a British Cohort", Royal Holloway University of London manuscript, 2003.

[3] Angrist, Joshua. D. and William. N. Evans . Schooling and Labor Market Consequences of the 1970 State Abortion Reforms. *Research in labor economics 18* (1999): 75-113.

[4] Aiyagari, S. Rao, Jeremy Greenwood, and Nezih Guner, 2000. "On the State of the Union." *Journal of Political Economy*, 109(2), 213-244.

[5] Anderson, Axel and Lones Smith. "Assortative Matching, Reputation, and the Beatles Break-Up." University of Michigan working paper, 2003.

[6] Anderson, Siwan. "Why dowry payments declined with modernization in Europe and are increasing in India", *Journal of Political Economy,* 111 (2), April 2003, p. 269-310.

[7] Bajari, Patrick and Jeremy T. Fox, "Collusion and Complementarities in an FCC Spectrum Auction," University of Chicago working paper, April 2005.

[8] Baker, Lynn, and Robert Emery. "When Every Relationship is Above Average." *Law and Human Behavior,* 17(4)*,* 1993. 439-450.

[9] Becker, Gary S. "A Theory of Marriage: Part I." *Journal of Political Economy* 81, no. 4 (1973): 813-46.

[27]Following Choo Siow (forthcoming 2006b) extension of CS, it should be straightforward, but useful, to extend this model to incorporate cohabitation.

[10] Becker, Gary S. “A Theory of Marriage: Part II.” *Journal of Political Economy* 82, no. 2, Part II (1974): S11-S26.

[11] Becker, Gary S. *A Treatise on the Family.* Cambridge, Mass.: Harvard University Press, 1991.

[12] Bergstrom, Theodore., and M. Bagnoli . “Courtship as a Waiting Game.” *Journal of Political Economy* 101, no. 1 (1993): 185-202.

[13] Bergstrom, Ted., and David Lam, ”The effects of cohort size on marriage markets in twentieth century Sweden” (September 24, 1989). Department of Economics, UCSB. http://repositories.cdlib.org/ucsbecon/bergstrom/1989

[14] Bisin, Alberto., Giorgio Topa and Thierry Verdier. “Religious Intermarriage and Socialization in the United States” *Journal of Political Economy*, Vol. 112 (3), June 2004, 615-664.

[15] Boulier, B. L. & Rosenzweig, M. R., ‘Schooling, search, and spouse selection: Testing economic theories of marriage and household behavior’, *Journal of Political Economy* 92, (1984): 712–732.

[16] Brien, Michael., Lee Lillard and Steven Stern. ”Cohabitation, Marriage, and Divorce in a Model of Match Quality,” *International Economic Review*, forthcoming.

[17] Burdett, K. & Coles, M. G. , ‘Marriage and class’, *Quarterly Journal of Economics* 112, (1997): 141–168.

[18] Casterline, John B., Lindy Williams, and Peter McDonald. “The Age Difference Between Spouses: Variations among Developing Countries.” *Population Studies* 40, no. 3 (1986): 353-374.

[19] Chiappori, Pierre-Andre., Murat Iriyigun and Yoram Weiss. “Marriage Markets, Divorce and Intra-Household Allocations”, University of Colorado working paper, 2005.

[20] Chiappori, Pierre-Andre and Phil Reny. “Matching to share risk”, University of Chicago working paper, 2005.

[21] Choo, Eugene and Aloysius Siow. “Who marries whom and why.” *Journal of Political Economy,* February 2006.

[22] Choo, Eugene and Aloysius Siow. “Marriage and cohabitation in Canada”, *Demography*, forthcoming 2006b.

[23] Compton, Janice and Robert A. Pollak. ”Why Are Power Couples Increasingly Concentrated in Large Metropolitan Areas,” NBER Working Papers 10918, 2004.

[24] Costa, Dora L. and Matthew E. Kahn. ”Power Couples: Changes In The Location Choice Of The College Educated,” *Quarterly Journal of Economics*, 2000, v115(4,Nov), 1287-1315.

[25] Dagsvik, John K. “Aggregation in Matching Markets.” *International Economic Review* 41, no. 1 (2000): 27-57.

[26] Fogli, A., Fernandez, R. and C. Olivetti. “Mothers and Sons: Preference Formation and Female Labor Force Dynamics” *Quarterly Journal of Economics*, 119 (4), 1249-1299, 2004.

[27] Fox, Jeremy T., "Estimating Matching Games With Endogenous Prices," University of Chicago working paper (2005),

[28] Giolito, Eugenio P. "A Search Model of Marriage with Differential Fecundity". Manuscript, Department of Economics, University of Maryland. (2003).

[29] Goldin, Claudia, and Lawrence Katz. "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions." *Journal of Political Economy* 110, no. 4 (2002): 730-70.

[30] Gould, Eric and M. Daniele Paserman. "Waiting for Mr. Right: Rising Inequality and Declining Marriage Rates," *Journal of Urban Economics*, 53 (2003), 257-281.

[31] Guner, N., Fernandez, R. and J. Knowles. "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality" *Quarterly Journal of Economic*, 120 (1), 273-344, 2005.

[32] Gretsky, Neil, Joseph Ostroy, and William Zame: "The Nonatomic Assignment Model," *Economic Theory*, 2(1), (1992): 102–127.

[33] Hamilton, Gillian and Aloysius Siow. "Class, Gender and Marriage", University of Toronto manuscript available at "www.economics.utoronto.ca/siow", 2000.

[34] Hansen, Lars and Thomas Sargent. *Recursive Models of Dynamic Linear Economies*. New York University manuscript, forthcoming.

[35] Hitsch, Günter, Ali Hortaçsu, and Dan Ariely, "What Makes You Click? An Empirical Analysis of Online Dating," 2005. University of Chicago working paper.

[36] Hotz, Joseph and Robert Miller. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *The Review of Economic Studies* 60, July 1993; 497-531.

[37] Iyigun, Murat F. and Randall P. Walsh. "Building the Collective Nest: A Collective Household Model with Competing Pre-Marital Investments and Spousal Matching," University of Colorado Working Paper No: 04-01, January 2004.

[38] Johansen, Rune, and John K. Dagsvik. "The Dynamics of a Behavioral Two-Sex Demographic Model." Statistics Norway, Research Department, 1999.

[39] Keeley, Michael C. "The Economics of Family Formation." *Economic Inquiry* 15(April), 1977: 238-250

[40] Kiernan, Kathleen. "European Perspectives on Union Formation." in *The Ties that Bind: Perspectives on Marriage and Cohabitation*, edited by Linda J. Waite, New York: Aldine de Gruyter, 2000.

[41] Lam, David. (1988), 'Marriage markets and assortative mating with household public goods: Theoretical results and empirical implications', *Journal of Human Resources* 23, 462–487.

[42] Lich-Tyler, Stephen. "Preference based matching in marriage markets", University of Michigan working paper, 2003.

[43] Lillard, Lee A. "Simultaneous Equations for Hazards — Marriage Duration and Fertility Timing." *Journal of Econometrics* 56, (1993): 189-217.

[44] Lundberg, Shelly and Robert Pollak. "Efficiency in Marriage" , *Review of Economics of the Household*, September 2003.

[45] Mahar, Heather, "Why are there so few prenupital agreements?" Harvard John M. Olin Discussion Paper Series: #436, 2003.

[46] Mare, D. (1991), 'Five decades of educational assortative mating', *American Sociological Review* 56, 15–32.

[47] McFadden, Daniel. "Conditional Logit Analysis of Qualitative Choice Behavior", in *Frontiers in Econometrics*, edited by P. Zarembka, New York: Academic Press, 1974

[48] Mortensen, Dale. "Matching: finding a partner for life or otherwise", *American Journal of Sociology* 94, supplement, 1988: S215-S240.

[49] Ni Bhrolchain, Maire. "Fexibility in the marriage market." *Population: An English Selction* 13, no.2 (2001): 9-48.

[50] Peters, Michael and Aloysius Siow. "Competing Pre-Marital Investments", *Journal of Political Economy*, June 2002, 592-608

[51] Pesendorfer, Martin and Philipp Schmidt-Dengler, "Identification and Estimation of Dynamic Games", September 2004, LSE working paper.

[52] Pollak, Robert. "Two-Sex Demographic Models." *Journal of Political Economy* 98, no. 2 (1990a): 399-420.

[53] ———— "Two-Sex Population Models and Classical Stable Population Theory." In *Convergent Issues in Genetics and Demography*, edited by Julian Adams and et. al., 317-33. New York, Oxford: Oxford University Press, 1990b.

[54] Pollard, John H. and Charlotte Hohn. "The Interaction between the Sexes." *Zeitschrift fur Bevolkerungswissenschaft* 19, no. 2 (1993-94): 203-08.

[55] Pollard, John H. "Modelling the Interaction between the Sexes." *Mathematical and Computer Modelling* 26, no. 6 (1997): 11-24.

[56] Qian, Zhenchao and Sam Preston. "Changes in American Marriage, 1972 to 1987: Availability and Forces of Attraction by Age and Education," *American Sociological Review,* 58, August (1993): 482-495.

[57] Qian, Zhenchao. "Changes in Assortative Mating: The Impact of Age and Education, 1970-1990." *Demography* 35, no. 3 (1998): 279-92.

[58] Roth, Alvin, and Marilda A. Oliveira Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis.* New York: Cambridge University Press, 1990.

[59] Rust, John. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55, no.5, (1987): 999-1033.

[60] Seitz, Shannon. "Employment and the Sex Ratio in a Two-Sided Model of Marriage." Queen's University manuscript, 1999.

[61] Shimer, Robert. & Smith, Lones. 'Assortative matching and search', *Econometrica* 68, (2000): 43–370.

[62] Siow, Aloysius. "Differential Fecundity, Markets and Gender Roles." *Journal of Political Economy* 106, No. 2 (1998): 334-354.

[63] Sorensen, Morten, "How Smart is Smart Money? An Empirical Two-Sided Matching Model of Venture Capital," December 2004. University of Chicago working paper.

[64] Svarer, Michael. "Determinants of Divorce in Denmark", Working Paper 2002-19, Department of Economics, University of Aarhus, 2002.

[65] Vella, Frank. and S. Collins. "The Value of Youth: Equalizing Age Differentials in Marriage." *Applied Economics* 22 (1990) 359-373.

[66] Waite, Linda, Ed. *The Ties that Bind: Perspectives on Marriage and Cohabitation.* Hawthorne, New York, Aldine De Gruyter, 2000.

[67] Weinstein, Neil. 1980. "Unrealistic Optimism About Future Life Events." *Journal of Personality and Soical Psychology 39*: 806-820.

[68] Weiss, Yoram and Robert J. Willis. "Match Quality, New Information, and Marital Dissolution", *Journal of Labor Economics* 15, no. 1, (1997): S293-S329.

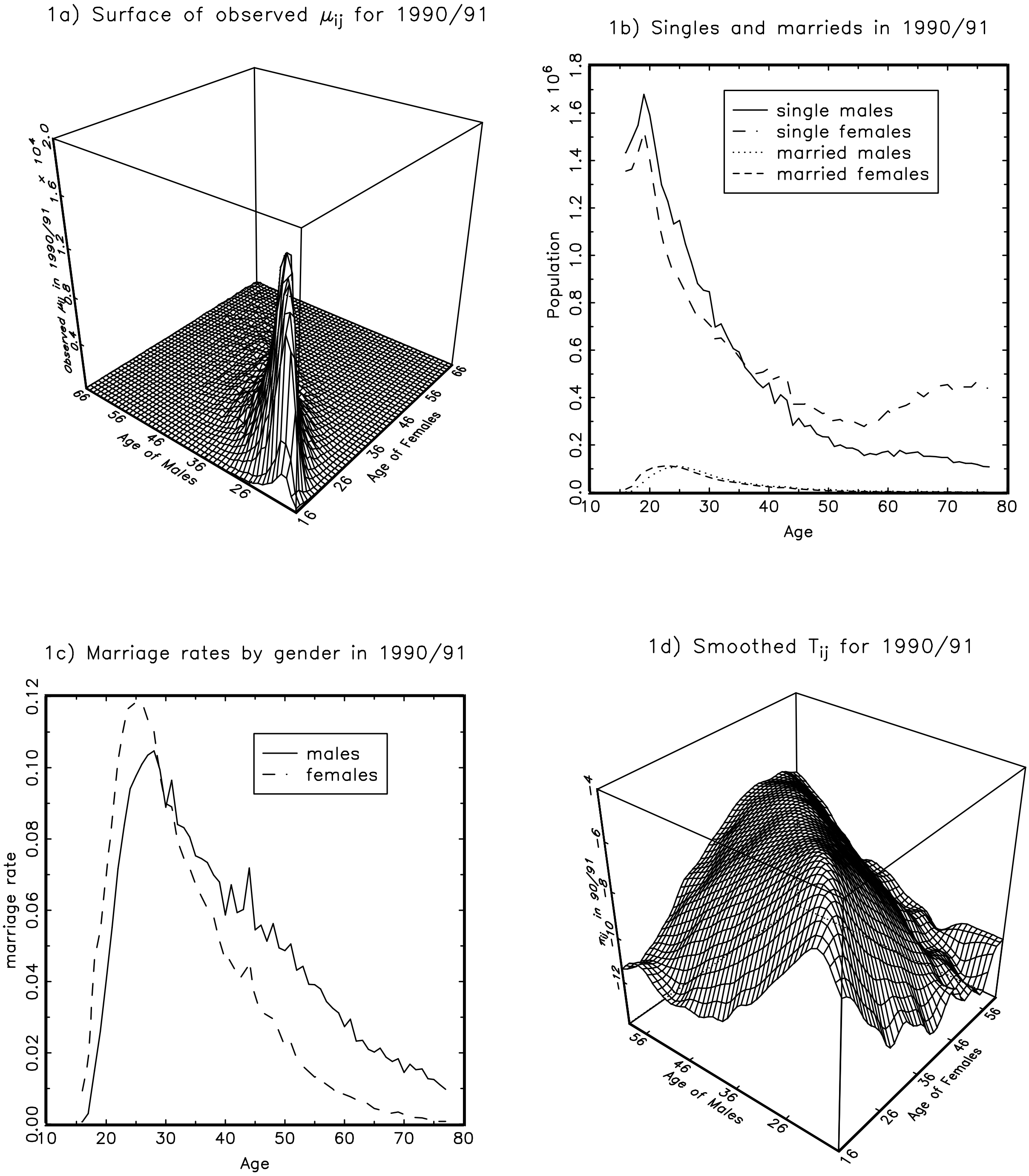
[69] Wong, Linda. "Structural estimation of marriage models." *Journal of Labor Economics* 27, no. 3 (2003): 699-728.

[70] Wong, Linda. "Aging in a two-sided matching model." SUNY Binghamton working paper (2005).

[71] Yatchew, Adonis. *Semiparametric Regression for the Applied Econometrician.* Cambridge: Cambridge University Press, 2003.

[72] Zhang, Junzen. " Do Men with HigherWages Marry Earlier or Later?" *Economics Letters* 49 (1995): 193-196.

Figure 1



1a) Surface of observed μ_{ij} for 1990/91
observed μ_{ij} in 1990/91 x 10^4
0.4
0.8
1.2
1.6
2.0
Age of Males
66
56
46
36
26
16
Age of Females
26
36
46
56
66
1b) Singles and marrieds in 1990/91
single males
single females
married males
married females
Population x 10^6
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
Age
10
20
30
40
50
60
70
80
1c) Marriage rates by gender in 1990/91
males
females
marriage rate
0.00
0.02
0.04
0.06
0.08
0.10
0.12
Age
10
20
30
40
50
60
70
80
1d) Smoothed T_{ij} for 1990/91
π_{ij} in 90/91
-12
-10
-8
-6
-4
Age of Males
56
46
36
26
16
Age of Females
26
36
46
56

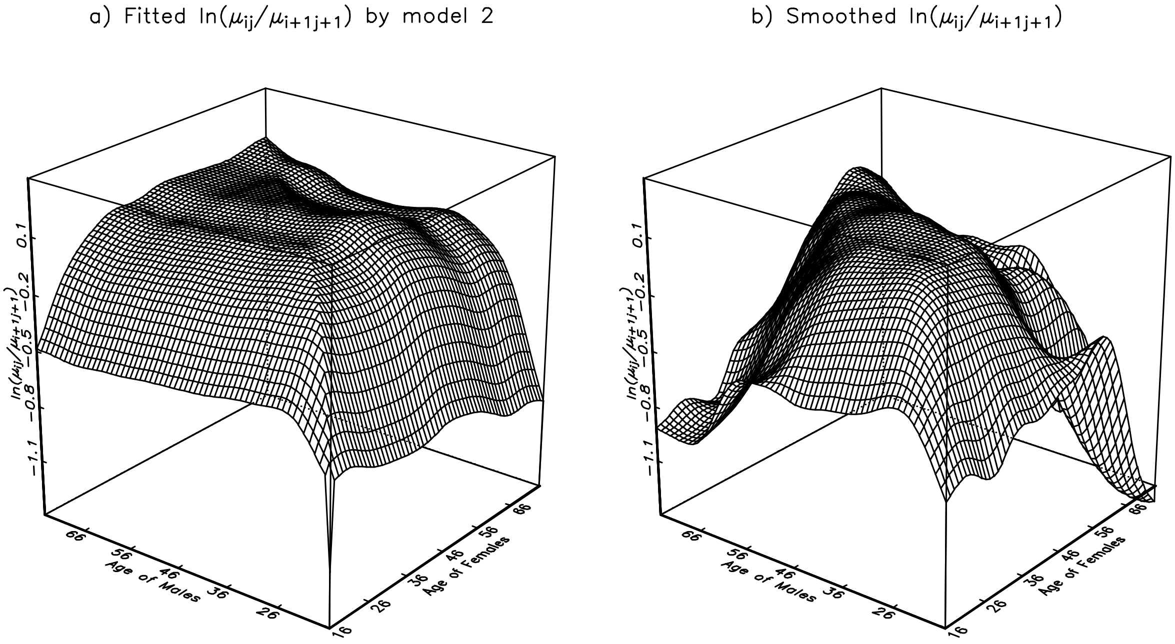


Figure 2, Fitted and Smoothed First Differenced

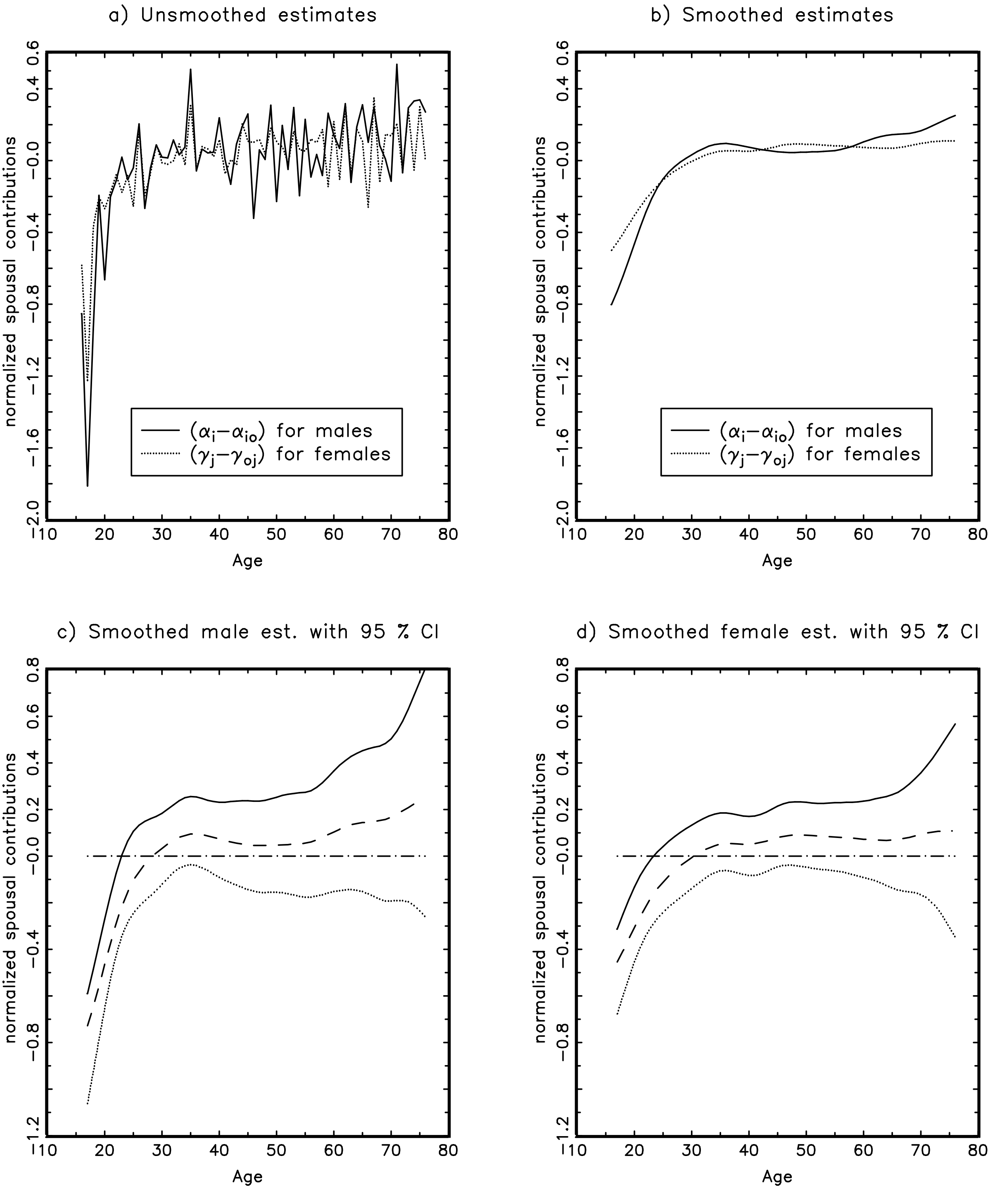


Figure 3, Estimated spousal contributions

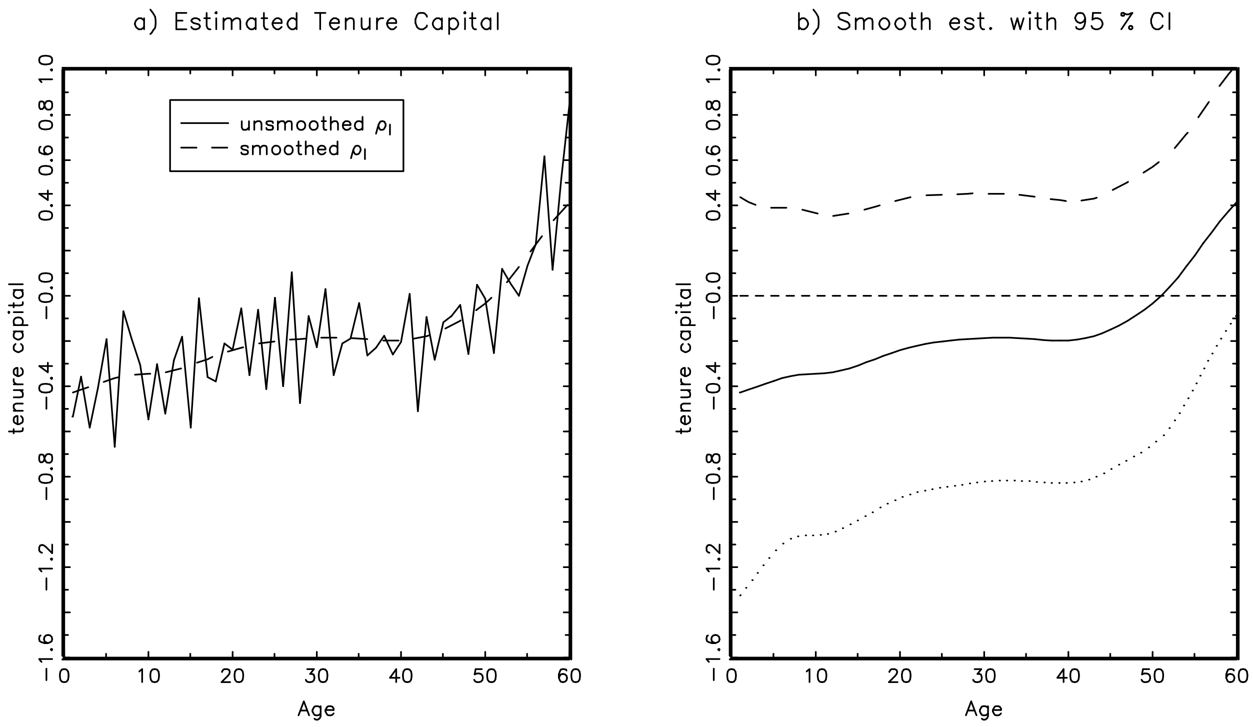
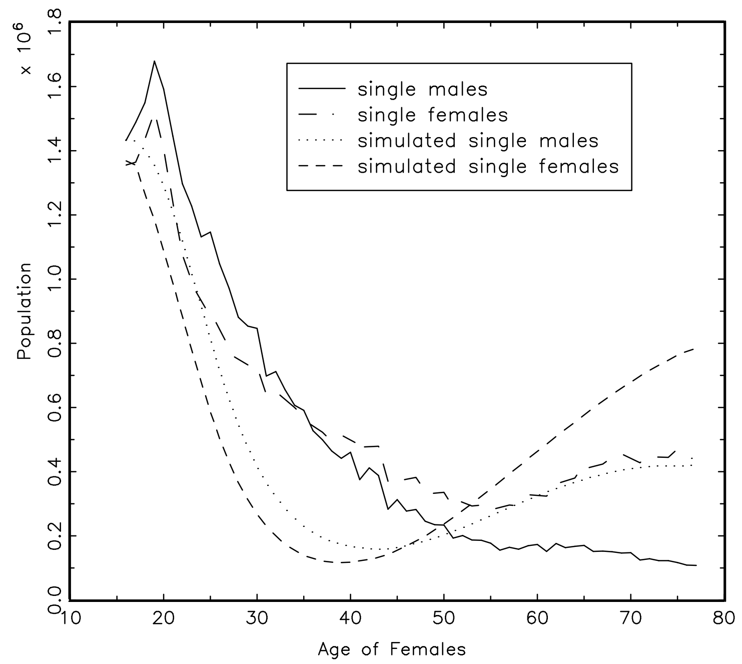


Figure 4, Estimated tenure capital

Figure 5, Simulated and observed unmarrieds in 1990/91



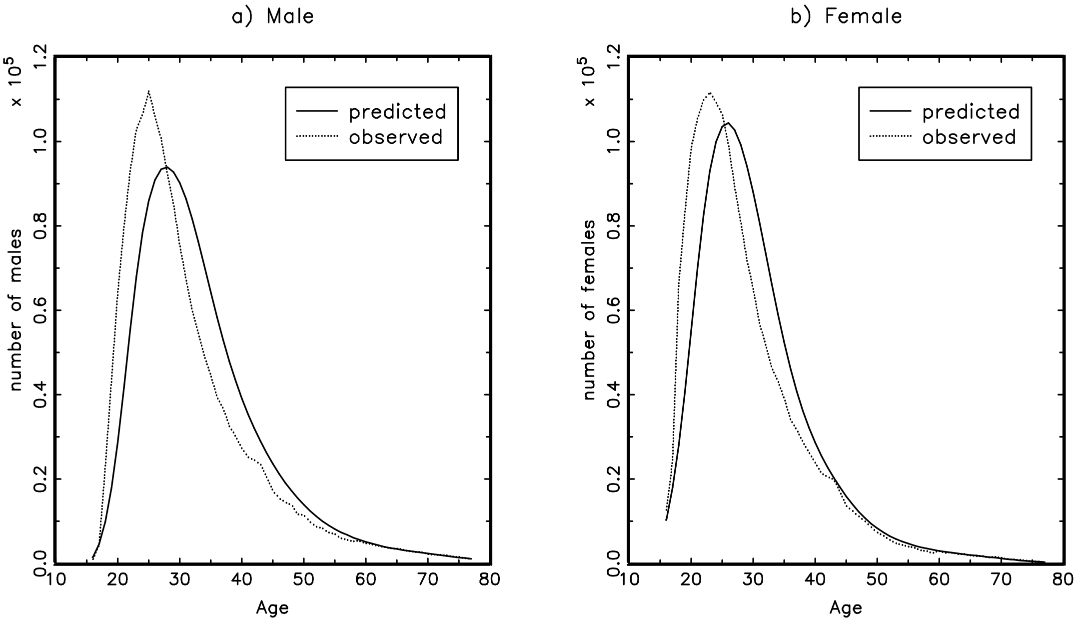


Figure 6, Observed and Predicted Marriages from model 2

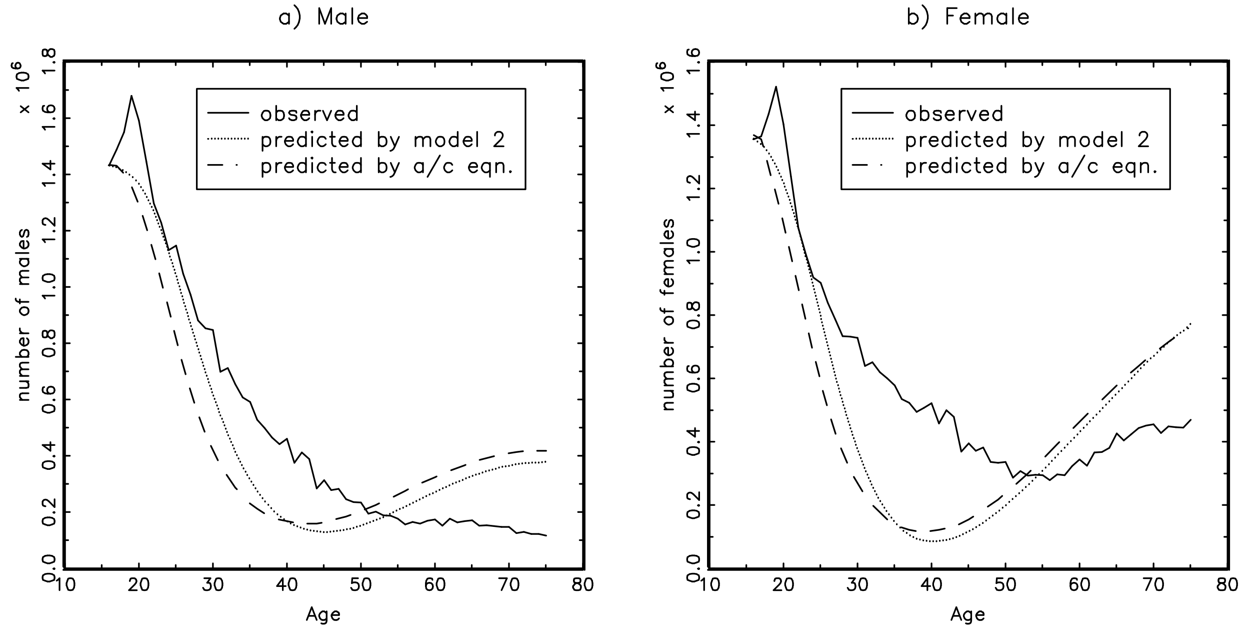


Figure 7, Observed and Predicted unmarrieds from model 2

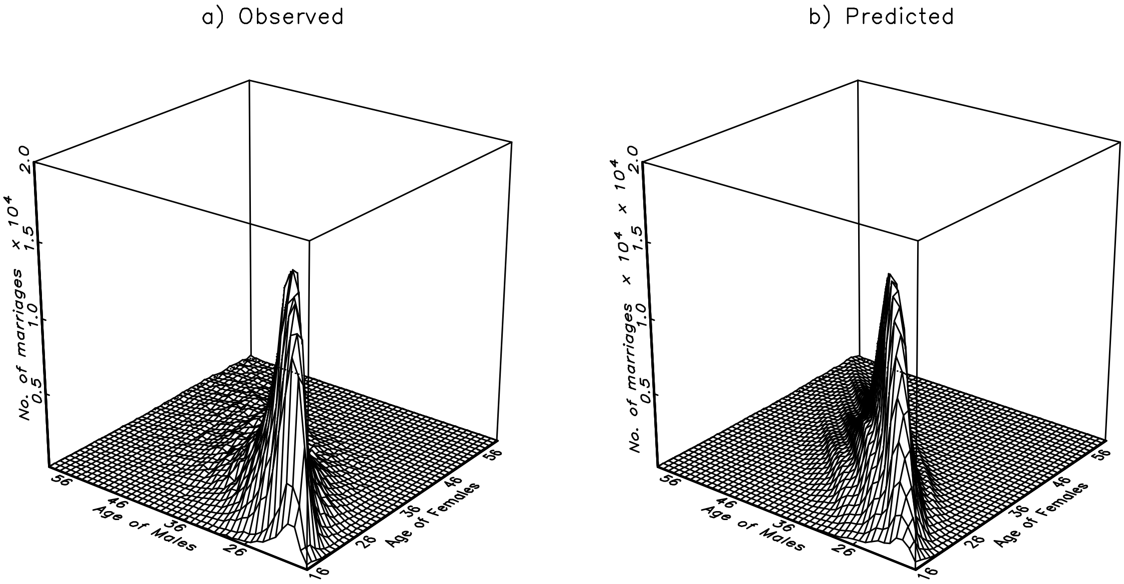


Figure 8, Observed and Predicted Marriage surface from model 2

8 Appendix

8.1 Generalizing the Divorce Hazard and Alternate Estimation Strategy

In this section, we consider the formulation of the model when the divorce hazard is not constant but depend on the types of the couple and duration of the match. Let δ_{ij}^d be the exogenous divorce hazard for an (i, j) couple in the $d'th$ year of the marriage. $d = 0$ denotes the first year of marriage and $\delta_{ij}^0 = 0$. The survival function, $S_{ij}^d = \prod_{l=0}^{d}(1 - \delta_{ij}^l)$, is the unconditional probability that this marriage will survive until year d. Divorce couples reenter the marriage market without penalty. The equation for the total gain from an (i, j) marriage relative to remaining single takes the form

$$
\begin{aligned}
T_{ij} &= \ln \frac{\mu_{ij}}{\sqrt{\mu_{i0}\mu_{0j}}} \\
&= \frac{1}{2}\sum_{k=0}^{z_{ij}} \beta^k S_{ij}^k \left(\pi_{ij}^k - \alpha_{i+k,0} - \gamma_{0,j+k} - 2c\right) + \frac{1}{2}\sum_{k=1}^{z_{ij}} \beta^k S_{ij}^k \ln \frac{m_{i+k+1} f_{j+k+1}}{\mu_{i+k+1,0}\mu_{0,j+k+1}} \qquad (24)
\end{aligned}
$$

The total gains to an (i, j) marriage relative to remaining single maintains a form similar to equation (12).

Our estimation strategy relies on taking a quasi-difference in total gains. This formulation requires that we admit less generality in the divorce hazard. Instead of having divorce hazard depend on the age of marriage of the couple (i, j), and duration d, let the divorce hazard only depend on the age of the married couple, that is

$$\delta_{ij}^d = \delta_{i+d,j+d}.$$

In other words, we assume that there is no duration dependence in divorce hazards, then the implied survival function take the form

$$S_{ij}^k = \prod_{l=0}^{k}(1 - \delta_{i+l,j+l}), \qquad \forall i, j, \text{ and } k$$

[28] Consider taking a quasi-first difference of Equation (24), we get

$$
\begin{aligned}
T_{ij} - \beta S_{ij}^1 T_{i+1,j+1} &= \frac{1}{2}\left[\pi_{ij}^0 - \alpha_{i0} - \gamma_{j0}\right] - \frac{1}{2}\beta S_{ij}^1 \left[2c - \ln \mathcal{P}_{i+1}\mathcal{Q}_{j+1}\right] \qquad (25) \\
&\quad + \frac{1}{2}\sum_{k=1}^{z_{ij}} \beta^k S_{ij}^k \left[\pi_{ij}^k - \pi_{i+1,j+1}^{k-1}\right]
\end{aligned}
$$

Equation (25) is an intertemporal arbitrage condition for marital matches similar to equation (14). $T_{ij} - \beta S_{ij}^1 T_{i+1,j+1}$ is the difference between the total gains to an ij marriage at time t and the appropriately discounted expected total gain to an $i + 1$, $j + 1$ match one period later. The first term on the right hand side of (25) is the difference in discounted within marriage payoffs. The second term reflects the difference in expected costs of forming an ij match in the current period rather than a year later. This difference consists of costs which are unique to the ij couple. The first component is the immediate

[28] This assumption permits the following recursion in survival functions

$$S_{ij}^k = S_{ij}^1 \cdot S_{i+1j+1}^{k-1}$$

cost of not being single. The second component is the discounted expected cost of not being able to be in the marriage market in the next period if the couple married this period. The cost is higher if the one period ahead marriage rates are higher.[29]

Let $\Delta\kappa_k = \kappa_k - \kappa_{k-1}$. Consider again decomposing the within marriage return π^k_{ij} according to 20. In the case where $\beta \neq 1$ and $\delta^d_{ij} = \delta$ for all i, j and d, substituting the spousal decomposition into the quasi-difference equation (25) gives us

$$T_{ij} - \beta S^1_{ij}\Big(c + \ln \frac{\mu_{i+1,j+1}}{m_{i+1} f_{j+1}} \sqrt{\mu_{i+1,0}\mu_{0,j+1}}\Big) =$$
$$\frac{1}{2}(\alpha_i - \alpha_{i0}) + \frac{1}{2}(\gamma_j - \gamma_{0j}) + \frac{1}{2}\sum_{k=1}^{z_{ij}} \beta^k S^k_{ij} \Delta\kappa_k + \psi_{ij} \tag{26}$$

Suppose we have data on divorce hazards δ^k_{ij}. Estimation of 26 can proceed by least squares if the orthogonality condition

$$\mathbf{E}_t(\psi_{ij} \mid \delta^k_{ij},\ \mu_{i+1,j+1},\ m_{i+1},\ f_{j+1}) = 0 \qquad \forall \quad i, j, k, t,$$

holds. If we are willing to set the discount factor β, the regression equation for this quasi-difference formulation is given by,

$$T_{ij} - \beta S^1_{ij} G_{ij} = D_i \widetilde{\alpha}_i + D_j \widetilde{\gamma}_j + \sum_{k=1}^{z_{ij}} R^k_{ij} \widetilde{\rho}_k + \psi^t_{ij} \tag{27}$$

where D_i and D_j are dummy variables that take a value 1 when the age of males and females are i and j respectively, and

$$G_{ij} = c + \ln \frac{\mu_{i+1,j+1}}{m_{i+1} f_{j+1}} \sqrt{\mu_{i+1,0}\mu_{0,j+1}}$$
$$R^k_{ij} = \begin{cases} \frac{1}{2}\beta^k S^k_{ij}\ , & \text{if } k \leq z_{ij} \\ 0\ , & \text{otherwise} \end{cases}$$

$\widetilde{\alpha}_i$ and $\widetilde{\gamma}_j$ estimate the current male and female relative contributions, $(\alpha_i - \alpha_{i0})$ and $(\gamma_j - \gamma_{0j})$ respectively. $\widetilde{\rho}_k$ estimates the increment in the value of marriage specific capital, $\Delta\kappa_k$. Estimating (27) identifies the current spousal contributions and changes in marriage specific capital. Again we cannot identify all the current relative spousal contributions. So we impose the following normalizations:

$$\sum_{i=1}^{z_{ij}} (\alpha_i - \alpha_{i0}) = 0$$
$$\sum_{j=1}^{z_{ij}} (\gamma_j - \gamma_{0j}) = 0$$

8.2 Appendix A: Standard errors and limiting distribution

We assume that the data $(\boldsymbol{\mu^t}, \boldsymbol{m^t}, \boldsymbol{f^t})$ are drawn from a stationary distribution.[30] Consider a universe where the unit of observations is the households. Let the $(Z^2 + 2Z) \times 1$ vector $\boldsymbol{\theta^t}$ denote the true

[29] c is a constant which can be ignored for the current discussion.

[30] Deriving the limiting distributions under the assumption of not stationarity is beyond the scope of this paper.

probability that we observe a particular type of household. For example, θ^t_{ij} denotes the probability that we observe a household of (i, j) married couple, θ^t_i the probability that we observe a household of a single i age male and so on. Given a random sample of size N on marriages at t and available men and women at t, $(\boldsymbol{\mu^t}, \boldsymbol{m^t}, \boldsymbol{f^t})$, the maximum likelihood estimator for θ^t_{ij} is $\widehat{\theta}^t_{ij} = \mu^t_{ij}/N$ where $N = \sum_{i=1}^Z \sum_{j=1}^Z \mu^t_{ij} + \sum_{i=1}^Z m^t_i + \sum_{j=1}^Z f^t_j$. The limiting distribution for this multinomial random vector is

$$\sqrt{N}(\widehat{\boldsymbol{\theta}}^t_N - \boldsymbol{\theta}^t) \rightsquigarrow N(0, \Omega),$$

where $\Omega^t = diag(\boldsymbol{\theta}^t) - \boldsymbol{\theta}^t \boldsymbol{\theta}'^t$. The quantity $\boldsymbol{T}^t = \chi(\boldsymbol{\theta}^t)$ defines the identified vector of marriage gains. The estimation strategies proposed use the maximum likelihood estimators for choice probabilities generating estimators of the form $\widehat{\boldsymbol{T}}^t = \chi(\widehat{\boldsymbol{\theta}}^t)$. Their limiting normal distribution can be derived by applying delta methods [ref Van Der Vaart Theorem 3.1]:

$$\sqrt{N}\Big(\chi(\widehat{\boldsymbol{\theta}}^t_N) - \chi(\boldsymbol{\theta}^t)\Big) \rightsquigarrow N\Big(0, (D_\theta \chi)\, \Omega\, (D_\theta \chi)'\Big),$$

where $D_\theta \chi$ is the matrix of first derivatives with respect to $\boldsymbol{\theta}^t$.

Let the $\Gamma\Gamma' = (D_\theta \chi)\, \Omega\, (D_\theta \chi)'$ denote the variance-covariance matrix of our estimator. In our proposed least squares estimation, we weight our estimates of $\boldsymbol{T}^t$ using Γ^{-1}. The linearity in Ω allows us to derive the following form for our weighting matrix,

$$\begin{aligned} \Gamma &= (D_\theta \chi)\Big(diag(\boldsymbol{\theta}^t) - \phi\, \boldsymbol{\theta}^t \boldsymbol{\theta}'^t\Big)(D_\theta \chi)\Psi'^{-1} \\ \text{where} \quad \Psi\Psi' &= (D_\theta \chi)\, diag(\boldsymbol{\theta}^t)\, (D_\theta \chi)', \\ \phi &= \frac{1 \pm \sqrt{1 + b'b}}{b'b} \\ \text{and} \quad b &= \Psi^{-1} (D_\theta \chi)\, \boldsymbol{\theta}^t. \end{aligned}$$

The matrix Ψ is square and invertible.

Table A1 : Parameter Estimates from various models (robust and efficient se)

	Model 1			Model 2			Model 3		
	Est.	std. error		Est.	std. error		Est.	std. error	
		robust	efficient		robust	efficient		robust	efficient
# of obs.	2512			2512			2512		
R^2	0.1489			0.1742			0.1743		
CON	-0.0340	0.0166	0.0073	0.1925	0.4348	0.3037	0.1141	0.5906	0.4443
$\widetilde{\alpha}_{17}$	-1.5266	0.2416	0.1859	-1.8121	0.2698	0.2078	-1.7708	0.3473	0.2592
$\widetilde{\alpha}_{18}$	-0.7436	0.1666	0.0911	-0.9434	0.1779	0.1046	-0.9033	0.2657	0.1968
$\widetilde{\alpha}_{19}$	-0.0495	0.1237	0.0630	-0.1926	0.1531	0.0783	-0.1538	0.2560	0.1809
$\widetilde{\alpha}_{20}$	-0.4742	0.0999	0.0777	-0.6638	0.1090	0.0944	-0.6258	0.2313	0.1755
$\widetilde{\alpha}_{21}$	-0.0765	0.0635	0.0574	-0.2026	0.0889	0.0738	-0.1661	0.2233	0.1684
$\widetilde{\alpha}_{22}$	-0.0039	0.0685	0.0579	-0.1125	0.0928	0.0746	-0.0773	0.2202	0.1640
$\widetilde{\alpha}_{23}$	0.0921	0.0916	0.0546	0.0188	0.1218	0.0719	0.0522	0.2135	0.1587
$\widetilde{\alpha}_{24}$	-0.0111	0.0785	0.0612	-0.1058	0.1105	0.0804	-0.0738	0.2035	0.1536
$\widetilde{\alpha}_{25}$	0.0701	0.0904	0.0671	-0.0397	0.1232	0.0901	-0.0098	0.1861	0.1413
$\widetilde{\alpha}_{26}$	0.2090	0.1354	0.0938	0.2041	0.1744	0.1186	0.2325	0.2228	0.1680
$\widetilde{\alpha}_{27}$	-0.1891	0.1502	0.0831	-0.2651	0.1983	0.1101	-0.2385	0.2194	0.1577
$\widetilde{\alpha}_{28}$	0.0563	0.0731	0.0527	-0.0399	0.1126	0.0807	-0.0151	0.1686	0.1255
$\widetilde{\alpha}_{29}$	0.0703	0.0525	0.0407	0.0876	0.0983	0.0652	0.1108	0.1612	0.1169
$\widetilde{\alpha}_{30}$	0.0857	0.0542	0.0422	0.0189	0.0866	0.0650	0.0410	0.1501	0.1140
$\widetilde{\alpha}_{31}$	0.0657	0.0470	0.0408	0.0159	0.0765	0.0624	0.0370	0.1412	0.1086
$\widetilde{\alpha}_{32}$	0.1518	0.0660	0.0477	0.1154	0.1022	0.0707	0.1349	0.1510	0.1078
$\widetilde{\alpha}_{33}$	-0.0031	0.0909	0.0539	0.0319	0.1111	0.0762	0.0495	0.1473	0.1044
$\widetilde{\alpha}_{34}$	0.1254	0.0606	0.0381	0.0747	0.1045	0.0625	0.0909	0.1353	0.0903
$\widetilde{\alpha}_{35}$	0.3749	0.0997	0.0623	0.5092	0.1435	0.0948	0.5242	0.1612	0.1073
$\widetilde{\alpha}_{36}$	0.0557	0.0740	0.0511	-0.0566	0.0992	0.0795	-0.0434	0.1230	0.0937
$\widetilde{\alpha}_{37}$	0.0618	0.0516	0.0443	0.0630	0.0775	0.0621	0.0746	0.0999	0.0763
$\widetilde{\alpha}_{38}$	0.0200	0.0524	0.0373	0.0419	0.0928	0.0628	0.0519	0.1093	0.0743
$\widetilde{\alpha}_{39}$	0.0703	0.0457	0.0342	0.0478	0.0821	0.0639	0.0566	0.0949	0.0728
$\widetilde{\alpha}_{40}$	0.2276	0.0927	0.0507	0.2391	0.1520	0.0858	0.2466	0.1596	0.0904
$\widetilde{\alpha}_{41}$	-0.0227	0.1119	0.0535	0.0071	0.1884	0.0933	0.0131	0.1930	0.0945
$\widetilde{\alpha}_{42}$	-0.0355	0.0777	0.0676	-0.1323	0.1306	0.0931	-0.1270	0.1314	0.0928
$\widetilde{\alpha}_{43}$	0.0963	0.0544	0.0519	0.0911	0.0890	0.0779	0.0947	0.0893	0.0773
$\widetilde{\alpha}_{44}$	0.1820	0.0736	0.0536	0.1901	0.1038	0.0822	0.1926	0.1041	0.0825
$\widetilde{\alpha}_{45}$	0.1716	0.1095	0.0785	0.2591	0.1760	0.1278	0.2602	0.1760	0.1283
$\widetilde{\alpha}_{46}$	-0.1827	0.1140	0.0648	-0.3217	0.1827	0.1074	-0.3221	0.1820	0.1082
$\widetilde{\alpha}_{47}$	0.0456	0.0654	0.0431	0.0589	0.1110	0.0814	0.0569	0.1124	0.0831
$\widetilde{\alpha}_{48}$	0.0853	0.0696	0.0613	0.0070	0.1066	0.0835	0.0037	0.1083	0.0874

continued on next page

Table A1: *continued from previous page*									
	Model 1			Model 2			Model 3		
	Est.	std. error		Est.	std. error		Est.	std. error	
		robust	efficient		robust	efficient		robust	efficient
$\widetilde{\alpha}_{49}$	0.1306	0.0938	0.0755	0.3087	0.1593	0.1190	0.3043	0.1610	0.1231
$\widetilde{\alpha}_{50}$	-0.0134	0.0954	0.0602	-0.2283	0.1709	0.0964	-0.2346	0.1723	0.1031
$\widetilde{\alpha}_{51}$	0.0582	0.0820	0.0501	0.1959	0.1352	0.0910	0.1882	0.1442	0.0999
$\widetilde{\alpha}_{52}$	0.0813	0.0853	0.0599	-0.0499	0.1285	0.1001	-0.0587	0.1351	0.1098
$\widetilde{\alpha}_{53}$	0.1564	0.1001	0.0641	0.2952	0.1750	0.1205	0.2853	0.1827	0.1290
$\widetilde{\alpha}_{54}$	-0.1042	0.1059	0.0735	-0.1954	0.1744	0.1387	-0.2069	0.1834	0.1463
$\widetilde{\alpha}_{55}$	0.1234	0.0817	0.0609	0.2302	0.1189	0.1046	0.2174	0.1398	0.1176
$\widetilde{\alpha}_{56}$	0.0120	0.0714	0.0568	-0.0906	0.1326	0.1100	-0.1051	0.1575	0.1242
$\widetilde{\alpha}_{57}$	0.0000	0.0707	0.0629	0.0349	0.1533	0.1030	0.0187	0.1795	0.1249
$\widetilde{\alpha}_{58}$	-0.0905	0.0721	0.0532	-0.0839	0.1611	0.0962	-0.1008	0.1881	0.1221
$\widetilde{\alpha}_{59}$	0.1335	0.0757	0.0602	0.2637	0.1432	0.1255	0.2454	0.1781	0.1487
$\widetilde{\alpha}_{60}$	0.0201	0.0913	0.0779	0.1402	0.1570	0.1538	0.1202	0.1884	0.1769
$\widetilde{\alpha}_{61}$	0.2120	0.1141	0.0758	0.0690	0.1437	0.1279	0.0473	0.1841	0.1569
$\widetilde{\alpha}_{62}$	0.0226	0.1439	0.0917	0.3160	0.1922	0.1477	0.2932	0.2267	0.1720
$\widetilde{\alpha}_{63}$	-0.1016	0.1141	0.0793	-0.1205	0.2111	0.1385	-0.1447	0.2491	0.1697
$\widetilde{\alpha}_{64}$	0.1167	0.1072	0.0772	0.1873	0.2154	0.1529	0.1602	0.2549	0.1824
$\widetilde{\alpha}_{65}$	0.0698	0.1070	0.0689	0.3097	0.1864	0.1445	0.2825	0.2302	0.1787
$\widetilde{\alpha}_{66}$	0.0289	0.1007	0.0758	0.1019	0.2033	0.1466	0.0735	0.2526	0.1837
$\widetilde{\alpha}_{67}$	0.0273	0.1242	0.0864	0.2955	0.1592	0.1380	0.2645	0.2259	0.1746
$\widetilde{\alpha}_{68}$	0.0148	0.1467	0.0860	0.0842	0.3043	0.1965	0.0531	0.3418	0.2281
$\widetilde{\alpha}_{69}$	0.0327	0.1353	0.0873	0.0082	0.3002	0.2096	-0.0250	0.3451	0.2428
$\widetilde{\alpha}_{70}$	0.0166	0.0875	0.0690	-0.1159	0.1496	0.1340	-0.1503	0.2270	0.1848
$\widetilde{\alpha}_{71}$	0.1424	0.1508	0.0941	0.5363	0.2616	0.1668	0.5010	0.3145	0.2093
$\widetilde{\alpha}_{72}$	-0.0310	0.1632	0.0875	-0.0680	0.3266	0.1688	-0.1026	0.3697	0.2138
$\widetilde{\alpha}_{73}$	0.1373	0.1426	0.0841	0.2946	0.2448	0.1583	0.2590	0.3056	0.2063
$\widetilde{\alpha}_{74}$	0.0130	0.1623	0.0961	0.3326	0.2563	0.1898	0.2964	0.3185	0.2324
$\widetilde{\alpha}_{75}$	0.2392	0.1478	0.1098	0.3374	0.3199	0.2418	0.3000	0.3722	0.2759
$\widetilde{\alpha}_{76}$	-0.0040	0.1790	0.1201	0.2699	0.3719	0.4908	0.2322	0.4167	0.5059
$\widetilde{\gamma}_{17}$	-1.0717	0.1159	0.1065	-1.2309	0.1173	0.1135	-1.1945	0.2388	0.2100
$\widetilde{\gamma}_{18}$	-0.2536	0.0945	0.0848	-0.3672	0.1016	0.0907	-0.3355	0.2257	0.1968
$\widetilde{\gamma}_{19}$	-0.1219	0.0893	0.0938	-0.2051	0.0967	0.1019	-0.1768	0.2195	0.2058
$\widetilde{\gamma}_{20}$	-0.1357	0.1007	0.0620	-0.2698	0.1083	0.0682	-0.2390	0.2221	0.1755
$\widetilde{\gamma}_{21}$	-0.0953	0.0993	0.0581	-0.1747	0.1106	0.0652	-0.1451	0.2300	0.1758
$\widetilde{\gamma}_{22}$	-0.0140	0.0720	0.0638	-0.0799	0.0821	0.0712	-0.0511	0.1992	0.1652
$\widetilde{\gamma}_{23}$	-0.1311	0.0817	0.0579	-0.1787	0.0902	0.0645	-0.1510	0.1950	0.1588
$\widetilde{\gamma}_{24}$	-0.0321	0.1147	0.0657	-0.0845	0.1293	0.0746	-0.0583	0.2026	0.1557
$\widetilde{\gamma}_{25}$	-0.1952	0.0840	0.0577	-0.2568	0.0923	0.0659	-0.2339	0.1849	0.1433

continued on next page

Table A1: *continued from previous page*									
	Model 1			Model 2			Model 3		
	Est.	std. error		Est.	std. error		Est.	std. error	
		robust	efficient		robust	efficient		robust	efficient
$\widetilde{\gamma}_{26}$	0.1567	0.0978	0.0613	0.1659	0.1179	0.0713	0.1897	0.1874	0.1455
$\widetilde{\gamma}_{27}$	-0.1563	0.1028	0.0710	-0.1966	0.1076	0.0775	-0.1760	0.1766	0.1390
$\widetilde{\gamma}_{28}$	-0.0263	0.0835	0.0638	-0.0709	0.0863	0.0673	-0.0525	0.1594	0.1246
$\widetilde{\gamma}_{29}$	0.0685	0.1136	0.0647	0.0880	0.1242	0.0709	0.1056	0.1725	0.1221
$\widetilde{\gamma}_{30}$	0.0233	0.0894	0.0702	-0.0134	0.1078	0.0833	0.0043	0.1583	0.1301
$\widetilde{\gamma}_{31}$	-0.0001	0.0921	0.0540	-0.0217	0.1101	0.0656	-0.0034	0.1530	0.1171
$\widetilde{\gamma}_{32}$	0.0121	0.0628	0.0385	-0.0034	0.0743	0.0474	0.0133	0.1312	0.1001
$\widetilde{\gamma}_{33}$	0.0509	0.0962	0.0559	0.0906	0.1196	0.0698	0.1053	0.1510	0.1064
$\widetilde{\gamma}_{34}$	-0.0015	0.0903	0.0488	-0.0250	0.1108	0.0606	-0.0107	0.1382	0.0962
$\widetilde{\gamma}_{35}$	0.2029	0.1069	0.0594	0.3126	0.1353	0.0776	0.3252	0.1514	0.1001
$\widetilde{\gamma}_{36}$	0.0203	0.0617	0.0378	-0.0492	0.0728	0.0503	-0.0369	0.1050	0.0807
$\widetilde{\gamma}_{37}$	0.0720	0.0594	0.0400	0.0771	0.0728	0.0524	0.0880	0.0986	0.0772
$\widetilde{\gamma}_{38}$	0.0363	0.0874	0.0507	0.0636	0.1031	0.0642	0.0724	0.1170	0.0793
$\widetilde{\gamma}_{39}$	0.0189	0.0677	0.0522	0.0206	0.0773	0.0592	0.0294	0.0932	0.0737
$\widetilde{\gamma}_{40}$	0.0903	0.0866	0.0564	0.1141	0.1152	0.0620	0.1218	0.1206	0.0728
$\widetilde{\gamma}_{41}$	-0.1074	0.0879	0.0635	-0.0745	0.1155	0.0837	-0.0684	0.1196	0.0864
$\widetilde{\gamma}_{42}$	0.0718	0.0705	0.0445	0.0057	0.0919	0.0619	0.0110	0.0950	0.0668
$\widetilde{\gamma}_{43}$	-0.0290	0.0808	0.0462	-0.0281	0.1048	0.0624	-0.0241	0.1046	0.0639
$\widetilde{\gamma}_{44}$	0.1977	0.0628	0.0378	0.2090	0.0827	0.0495	0.2122	0.0837	0.0510
$\widetilde{\gamma}_{45}$	0.0340	0.0828	0.0406	0.1062	0.1134	0.0625	0.1081	0.1132	0.0627
$\widetilde{\gamma}_{46}$	0.2110	0.0895	0.0464	0.1001	0.1294	0.0675	0.1006	0.1295	0.0675
$\widetilde{\gamma}_{47}$	0.0971	0.0836	0.0654	0.1138	0.1144	0.0817	0.1134	0.1158	0.0827
$\widetilde{\gamma}_{48}$	0.0825	0.0880	0.0579	0.0331	0.0866	0.0656	0.0315	0.0884	0.0667
$\widetilde{\gamma}_{49}$	0.0498	0.0759	0.0542	0.1824	0.1054	0.0690	0.1797	0.1093	0.0721
$\widetilde{\gamma}_{50}$	0.2580	0.0976	0.0635	0.1026	0.1191	0.0678	0.0987	0.1237	0.0726
$\widetilde{\gamma}_{51}$	-0.0233	0.1105	0.0761	0.0774	0.1535	0.0733	0.0720	0.1608	0.0809
$\widetilde{\gamma}_{52}$	0.0816	0.1216	0.0681	-0.0194	0.1179	0.0675	-0.0260	0.1275	0.0783
$\widetilde{\gamma}_{53}$	0.0634	0.0899	0.0611	0.1646	0.0991	0.0649	0.1567	0.1113	0.0800
$\widetilde{\gamma}_{54}$	0.1353	0.0856	0.0620	0.0587	0.1138	0.0767	0.0493	0.1300	0.0928
$\widetilde{\gamma}_{55}$	-0.0161	0.1099	0.0687	0.0541	0.1139	0.0735	0.0438	0.1351	0.0946
$\widetilde{\gamma}_{56}$	0.2059	0.1134	0.0809	0.1216	0.0907	0.0754	0.1101	0.1212	0.1005
$\widetilde{\gamma}_{57}$	0.0815	0.1053	0.0765	0.0997	0.1059	0.0740	0.0870	0.1357	0.1034
$\widetilde{\gamma}_{58}$	0.1697	0.1355	0.0754	0.1729	0.1269	0.0713	0.1583	0.1562	0.1072
$\widetilde{\gamma}_{59}$	-0.2335	0.1259	0.0799	-0.1478	0.1141	0.0776	-0.1636	0.1507	0.1161
$\widetilde{\gamma}_{60}$	0.1437	0.1144	0.0666	0.2166	0.1404	0.0878	0.1999	0.1727	0.1269
$\widetilde{\gamma}_{61}$	0.0019	0.1205	0.0704	-0.1059	0.1037	0.0768	-0.1238	0.1540	0.1240
$\widetilde{\gamma}_{62}$	0.1184	0.1344	0.0831	0.3089	0.1365	0.0872	0.2894	0.1829	0.1372
continued on next page									

Table A1: *continued from previous page*									
	Model 1			Model 2			Model 3		
	Est.	std. error		Est.	std. error		Est.	std. error	
		robust	efficient		robust	efficient		robust	efficient
$\widetilde{\gamma}_{63}$	-0.0490	0.1327	0.0706	-0.0744	0.1508	0.0854	-0.0951	0.2010	0.1395
$\widetilde{\gamma}_{64}$	0.1456	0.1658	0.1069	0.1734	0.1552	0.1004	0.1529	0.2108	0.1545
$\widetilde{\gamma}_{65}$	-0.0675	0.1320	0.1125	0.1034	0.1314	0.0976	0.0804	0.1957	0.1567
$\widetilde{\gamma}_{66}$	-0.2796	0.1583	0.1127	-0.2625	0.1448	0.0996	-0.2868	0.2086	0.1602
$\widetilde{\gamma}_{67}$	0.1935	0.1261	0.1266	0.3523	0.1177	0.1031	0.3279	0.1892	0.1711
$\widetilde{\gamma}_{68}$	-0.1163	0.1070	0.0885	-0.1235	0.1798	0.1205	-0.1502	0.2373	0.1809
$\widetilde{\gamma}_{69}$	0.2022	0.1103	0.1050	0.1446	0.1969	0.1403	0.1177	0.2557	0.1984
$\widetilde{\gamma}_{70}$	0.2921	0.1399	0.1223	0.1385	0.1157	0.1079	0.1106	0.2054	0.1785
$\widetilde{\gamma}_{71}$	-0.0716	0.1252	0.1434	0.2068	0.2082	0.1249	0.1776	0.2685	0.1949
$\widetilde{\gamma}_{72}$	0.0151	0.1864	0.1389	-0.0598	0.3063	0.1302	-0.0911	0.3430	0.2013
$\widetilde{\gamma}_{73}$	0.2104	0.1722	0.1547	0.2852	0.2318	0.1399	0.2532	0.2959	0.2095
$\widetilde{\gamma}_{74}$	-0.2586	0.2181	0.1674	-0.0541	0.2258	0.1664	-0.0868	0.2881	0.2276
$\widetilde{\gamma}_{75}$	0.2408	0.2985	0.2094	0.2980	0.2505	0.2410	0.2652	0.3119	0.2873
$\widetilde{\gamma}_{76}$	-0.1605	0.3281	0.2804	-0.0011	0.1866	0.3942	-0.0349	0.2629	0.4240
$\widetilde{\rho}_{1}$				-0.5342	0.6446	0.6139	-0.3842	1.0000	0.8614
$\widetilde{\rho}_{2}$				-0.3578	0.5863	0.4114	-0.2102	0.9658	0.7404
$\widetilde{\rho}_{3}$				-0.5835	0.5527	0.3801	-0.4371	0.9326	0.7187
$\widetilde{\rho}_{4}$				-0.4095	0.5378	0.3614	-0.2647	0.9223	0.7080
$\widetilde{\rho}_{5}$				-0.1905	0.5817	0.3629	-0.0476	0.9260	0.7038
$\widetilde{\rho}_{6}$				-0.6689	0.5364	0.3642	-0.5286	0.9077	0.6958
$\widetilde{\rho}_{7}$				-0.0672	0.4897	0.3474	0.0692	0.8743	0.6753
$\widetilde{\rho}_{8}$				-0.1892	0.5397	0.3730	-0.0541	0.8943	0.6829
$\widetilde{\rho}_{9}$				-0.3073	0.5451	0.3685	-0.1745	0.8779	0.6730
$\widetilde{\rho}_{10}$				-0.5467	0.4951	0.3539	-0.4185	0.8409	0.6641
$\widetilde{\rho}_{11}$				-0.3026	0.5065	0.3522	-0.1754	0.8472	0.6510
$\widetilde{\rho}_{12}$				-0.5214	0.4966	0.3482	-0.3949	0.8256	0.6405
$\widetilde{\rho}_{13}$				-0.2860	0.5083	0.3500	-0.1653	0.8338	0.6313
$\widetilde{\rho}_{14}$				-0.1803	0.5039	0.3405	-0.0602	0.8141	0.6176
$\widetilde{\rho}_{15}$				-0.5836	0.5028	0.3449	-0.4654	0.7917	0.6088
$\widetilde{\rho}_{16}$				-0.0096	0.4855	0.3355	0.1045	0.7786	0.5950
$\widetilde{\rho}_{17}$				-0.3603	0.4833	0.3408	-0.2482	0.7626	0.5905
$\widetilde{\rho}_{18}$				-0.3796	0.4778	0.3326	-0.2686	0.7572	0.5793
$\widetilde{\rho}_{19}$				-0.2104	0.4829	0.3262	-0.1021	0.7494	0.5663
$\widetilde{\rho}_{20}$				-0.2393	0.4748	0.3277	-0.1342	0.7352	0.5549
$\widetilde{\rho}_{21}$				-0.0562	0.4715	0.3290	0.0471	0.7215	0.5440
$\widetilde{\rho}_{22}$				-0.3508	0.4718	0.3268	-0.2502	0.7047	0.5333
$\widetilde{\rho}_{23}$				-0.0617	0.4730	0.3304	0.0364	0.6913	0.5277

continued on next page

Table A1: *continued from previous page*									
	Model 1			Model 2			Model 3		
	Est.	std. error		Est.	std. error		Est.	std. error	
		robust	efficient		robust	efficient		robust	efficient
$\widetilde{\rho}_{24}$				-0.4125	0.4726	0.3251	-0.3169	0.6813	0.5173
$\widetilde{\rho}_{25}$				-0.0087	0.4734	0.3230	0.0848	0.6658	0.5072
$\widetilde{\rho}_{26}$				-0.4001	0.4684	0.3223	-0.3090	0.6640	0.5003
$\widetilde{\rho}_{27}$				0.1053	0.4704	0.3206	0.1938	0.6470	0.4898
$\widetilde{\rho}_{28}$				-0.4752	0.4618	0.3215	-0.3897	0.6373	0.4828
$\widetilde{\rho}_{29}$				-0.0893	0.4597	0.3194	-0.0062	0.6204	0.4687
$\widetilde{\rho}_{30}$				-0.2271	0.4553	0.3193	-0.1465	0.6152	0.4612
$\widetilde{\rho}_{31}$				0.0299	0.4667	0.3171	0.1078	0.6082	0.4516
$\widetilde{\rho}_{32}$				-0.3505	0.4629	0.3171	-0.2754	0.5999	0.4447
$\widetilde{\rho}_{33}$				-0.2116	0.4538	0.3136	-0.1392	0.5785	0.4334
$\widetilde{\rho}_{34}$				-0.1900	0.4498	0.3070	-0.1202	0.5720	0.4230
$\widetilde{\rho}_{35}$				-0.0318	0.4502	0.3244	0.0352	0.5620	0.4233
$\widetilde{\rho}_{36}$				-0.2630	0.4616	0.3149	-0.1983	0.5542	0.4119
$\widetilde{\rho}_{37}$				-0.2304	0.4554	0.3192	-0.1687	0.5429	0.4005
$\widetilde{\rho}_{38}$				-0.1758	0.4449	0.3155	-0.1171	0.5268	0.3899
$\widetilde{\rho}_{39}$				-0.2586	0.4490	0.3099	-0.2020	0.5219	0.3805
$\widetilde{\rho}_{40}$				-0.2043	0.4410	0.3099	-0.1510	0.5080	0.3728
$\widetilde{\rho}_{41}$				0.0089	0.4544	0.3095	0.0593	0.5102	0.3659
$\widetilde{\rho}_{42}$				-0.5104	0.4511	0.3115	-0.4628	0.5050	0.3626
$\widetilde{\rho}_{43}$				-0.0943	0.4413	0.3075	-0.0495	0.4883	0.3510
$\widetilde{\rho}_{44}$				-0.2836	0.4622	0.3146	-0.2412	0.4993	0.3513
$\widetilde{\rho}_{45}$				-0.1176	0.4377	0.3053	-0.0786	0.4696	0.3366
$\widetilde{\rho}_{46}$				-0.0891	0.4371	0.3067	-0.0531	0.4650	0.3323
$\widetilde{\rho}_{47}$				-0.0405	0.4346	0.3062	-0.0062	0.4577	0.3280
$\widetilde{\rho}_{48}$				-0.2577	0.4498	0.2711	-0.2253	0.4691	0.2940
$\widetilde{\rho}_{49}$				0.0485	0.5020	0.3480	0.0783	0.5161	0.3625
$\widetilde{\rho}_{50}$				-0.0123	0.4665	0.3194	0.0140	0.4819	0.3306
$\widetilde{\rho}_{51}$				-0.2523	0.4384	0.3103	-0.2298	0.4475	0.3178
$\widetilde{\rho}_{52}$				0.1193	0.4316	0.3031	0.1403	0.4378	0.3099
$\widetilde{\rho}_{53}$				0.0531	0.4136	0.2980	0.0702	0.4175	0.3018
$\widetilde{\rho}_{54}$				0.0007	0.4249	0.2995	0.0154	0.4269	0.3013
$\widetilde{\rho}_{55}$				0.1348	0.3928	0.2942	0.1469	0.3931	0.2952
$\widetilde{\rho}_{56}$				0.2276	0.3907	0.2871	0.2378	0.3891	0.2867
$\widetilde{\rho}_{57}$				0.6175	0.3807	0.2797	0.6255	0.3803	0.2796
$\widetilde{\rho}_{58}$				0.1139	0.3941	0.2747	0.1218	0.3938	0.2731
$\widetilde{\rho}_{59}$				0.5177	0.3843	0.2626	0.5221	0.3840	0.2622
$\widetilde{\rho}_{60}$				0.8460	0.3212	0.2858	0.8466	0.3212	0.2858

continued on next page

Table A1: *continued from previous page*									
	Model 1			Model 2			Model 3		
	Est.	std. error		Est.	std. error		Est.	std. error	
		robust	efficient		robust	efficient		robust	efficient
$(\ln(i) - \ln(j))^2$							-0.8922	4.2305	3.6447
$(\ln(i) - \ln(j))^3$							0.2311	6.3621	5.0157
$(\ln(i) - \ln(j))^4$							4.6443	23.6322	22.7090