Dynamic marriage matching: An Empirical Framework*

Eugene Choo

University of Calgary

Draft: 21 March 2012

Abstract

The paper proposes a dynamic version of the frictionless Becker-Shapley-Shubik marriage matching model with transferable utility. Its primary objective is to develop a tractable model that rationalizes the marriage distribution of 'who marries whom' by age. This behavioral dynamic model rationalizes a new marriage matching function. An empirical methodology that relies on the equilibrium outcomes of the model identifies the marital preferences over spouses. This framework also allows the inverse problem of computing the vector of aggregate marriages given a new distribution of available single individuals and estimated preferences to be solved. The solution to this inverse problem has been shown to exist under mild conditions. This paper also develops a simple test of the model's empirical validity. Using aggregate data of new marriages and available single men and women in the US in 1970, 1980 and 1990, I investigate the changes in the gains to marriage over this period.

*I thank Shashi Khatri, Harry Paarsch, Joanne Roberts and Matt Shum for their insightful comments. I also thank seminar participants at Singapore Management University and National University of Singapore (NUS). Part of this research was done while on sabbatical at the Economics Department at NUS, I thank the department there for their hospitality and support. I also thank SSHRC grants # 410-2010-2584 and 410-2006-1254 for financial support.

1 Introduction

The marriage distribution of 'who marries whom' by age shows many well known empirical regularities. There is strong assortative matching by age with men marrying slightly younger women. As the number of available single men and women falls with age, so do the marriage rates. While these empirical regularities are not static and have changed over time and differ across countries, these qualitative features have generally remain consistent. The timing of marriage has a significant effect on the formation and organization of families, including the timing of childbirth, the division of home production, etc.

The paper proposes a dynamic version of the frictionless Becker-Shapley-Shubik marriage matching model using transferable utility. Its primary objective is to develop a tractable dynamic model that rationalizes the marriage distribution of 'who marries whom' by age together with its empirical regularities. It also extends the static frictionless marriage matching framework of Choo and Siow (2006) into an overlapping generations framework. The behavioral dynamic model rationalizes a new marriage matching function. I develop an empirical methodology to identify preferences over spouse from marriage that relies on the equilibrium outcomes of the model. It provides an economic interpretation of these estimated parameters.[1] This framework also allows the inverse problem of computing the vector of aggregate marriage given a new distribution of available single individuals and estimated preferences to be solved. The proposed type of transferable utility matching model that has been shown to be equivalent to an optimal transportation linear programming problem. The solution to this type of problems has been shown to exist under mild conditions. This paper also develops a simple test of the model's empirical validity. Using aggregate data of new marriages and available single men and women in the US in the 1970 and 1980, I demonstrate its application by looking at the the changes in the gains in marriage over this decade and compare the results with those obtained from a static model.

[1] In a stationary environment, these estimated parameters are invariant to the confounding effects of changes in the number of available single individuals over the life-cycle.

Over the two decades from 1970 to 1990, there has been a well documented fall in the marriage rates in the US. Part of this decline can be explained by socio-political changes that affected the institution of marriage. Changes like the national legalization of abortion following the Supreme Court ruling on Roe versus Wade has been argued to lower the gains to marriage. The empirical methodology developed in this paper allows the total gains to marriage to be identified. This is the present discounted net present value from marriage today relative to the present discounted per period utility from remaining single (forever). The empirical analysis shows that the dynamic component of the gains to marriage is a large component of the total gains from marriage. This is especially true since most marriage occur when individuals are young when there are still many future opportunities of participating in the marriage market as the individuals age. The decision to marry early suggests that the implied present discounted relative returns from locking into marriage early is high. When analyzing the change in the gains to marriage over these two decades, I show that ignoring the dynamic component of marriage gains severely understate the decline in the gains to marriage among the young.

The first empirical implementation of the static Becker-Shapley-Shubik marriage matching model with additively separable utilities in a discrete choice framework was proposed in Choo and Siow (2006). This paper maintains many of the minimal a priori assumptions of the static marriage matching model of Choo and Siow (2006). Methodologically, the formulation of the model in this paper uses the dynamic discrete choice framework of Rust (1987). The joint payoff to a match depends on the ages of the couple. Each cohort of single males and females enters the marriage market at age zero. At each age, a single individual faced with the marital returns associated with his or her age decides whether to marry or remain single. In this dynamic environment, the single individual understand that his or her type changes over time as he or she ages. The agents are rational and have an expectation of the marriage opportunities in the future as they age.

An earlier paper of Choo and Siow (2005) also shares a similar objective of attempting to model the bivariate marriage distribution by age and uses the same building block of the dynamic discrete choice framework of Rust (1987). While it shares this similarities, these papers differ in many ways. Choo and Siow (2005) proposes a general equilibrium framework that allows different theories of marital and home production to be tested. Choo and Siow (2005) embeds a marriage matching framework into a set of endogenous population accounting equations. These different theories are tested by putting restrictions on a linearized per period gains from marriage. Choo and Siow (2005) provides a nice and intuitive approximation to the per period gains from marriage.[2] The general equilibrium structure of the model in Choo and Siow (2005) made it impossible to solve the model. Instead of solving the model, Choo and Siow (2005) took a different approach where a linearized structure was placed on the per period utilities from marriage. This linearization allowed us to empirically approximate the benefit of delaying marriage for one period versus marrying today using the growth rate of marriages. This approximation become the basis on which different theories of marital and home production were tested.

The dynamic model in this paper takes a more modest partial equilibrium approach. It takes the vector of available single individuals at the beginning of each period as given. I do not model how this vector of single individuals evolve dynamically and how it is affected by mortality, migration and marriage. This paper also proposes a very different representation of the dynamic problem that is empirically tractable. This new representation allows me to solve the model and derive the implied closed form marriage matching function. This new marriage matching function is the dynamic analogue of the one proposed in Choo and Siow (2006). I also propose a new empirical methodology that relies on the equilibrium outcomes of the model to identify the primitives of the model. Unlike Choo and Siow (2005), no structure is placed on the per period marital returns. This paper also focuses on identification of net present discounted utility from a match and the inverse problem associated with the application of the marriage matching function.

[2]The survey paper of Siow (2008) also provides a two period detailed exposition of the dynamic model of Choo and Siow (2005).

Related Literature:

There is a growing body of empirical papers on marriage matching. Choo and Siow (2006) proposed an equilibrium model of static marriage matching using the the discrete choice framework. Browning, Chiappori and Weiss (2007) and Chiappori, Salaniè and Weiss (2010) provides a stable matching characterization for preference utility that maintain the additive separability structure first introduced in Choo and Siow (2006). They showed that the additively separable structure reduces the complexity of stable matching into a set of simple inequalities easily satisfied by the probabilistic discrete choice framework of Choo and Siow (2006). A number of papers have proposed generalizations to the Choo and Siow's empirical framework. Chiappori, Salaniè and Weiss (2010) and Galichon and Salanié (2011) both have proposed distributional generalizations to the idiosyncratic shocks allowing for heterogeneity. The former paper uses this more general characterization to investigate the marital college premium. Focusing on the social surplus function as the basis for their empirical application, Galichon and Salanié(2010) also propose parametric approach that allows matching across many observable attributes. These papers maintain a static characterization of marriage matching which the current paper generalizes.

Decker, Lieb, McCann and Stephens (2010) provides a simple and elegant test for the Choo and Siow model. Their test exploits symmetry restrictions on the cross type marriage elasticity matrix. In words, the symmetry restriction requires that the elasticity of type i single men to the supply of type j women be equal to the elasticity of the type j single women with respect the supply of type i men. This restriction is reminiscent of the Independence of Irrelevant Alternatives (IIA) property brought about by the i.i.d. additive utility error imposed by the discrete choice structure (see McFadden (1973) and Debreu (1960)). The exact cost of this restriction in the context marriage matching model maintaining the Choo and Siow structure remains to be seen. A large body of literature in empirical Industrial Organization is devoted to overcoming the IIA properties of the discrete choice models.

Fox (2009) focuses on identifying preferences in transferable utility models applied to individual-level data. He proposes a maximum-score estimator based on a "pairwise stability" requirement. When the matching markets gets large, these pairwise stability conditions become computationally infeasible. Fox shows that as long as the "rank-order condition" holds, only a subset of the inequalities need to be used in the estimation.

There is a few papers that look at the empirical implications of the non-transferable utility models on individual level matchings. Most notable is Echenique, Lee, Shum and Yenmez (2011) which is the first paper to derived the empirical implications of stable two-sided matching. The authors develop a revealed preference theory for stable matching and propose a non-parametric test for stability. They show that transferable utility matching theory is empirically nested in non-transferable utility matching theory.

There is also a growing number of empirical papers that investigate the nature of marriage matching preferences through field experiments. Dugar, Bharttacharya and Reiley (2011) conduct an experiment to analyze how single men and women are willing to trade social status in India in the form of caste in the presence of strong economic incentives using a reputable (real world) Bengali arranged marriage market. The authors placed newspaper advertisements of potential grooms that systematically vary their caste and income and focused on responses of higher-caste females to lower-caste males.[3] The authors provide strong empirical evidence suggesting that despite strong caste-based discrimination, higher-status females are willing to trade caste-status for an increase in the advertised income of lower-status males.[4]

A closely related paper is Hitsch, Hortaçsu and Ariely (2010), which focus on identifying preferences separately from the matching process. Employing a dataset from an online dating service, they estimate a rich specification of preference over spousal physical and socio-economic characteristics. Using the estimated preferences, they simulate the men- and women-optimal matchings using the Gale-Shapley's deferred acceptance algorithm. The paper goes on to compare these optimal matchings to the actual matches observed in the online dating dataset.

[3]Higher-caste females lose their caste-status if they marry a lower-caste male.

[4]Using a major online dating website in South Korea, Lee and Niederle (2011) conducted an experiment to see how preference signaling in the form of a virtual rose can help chances in online dating.

2 A Dynamic Matching Model

2.1 *Preview of Results*

The equilibrium dynamic marriage matching model in this paper delivers a new marriage matching function given by equation (1) below. A *marriage matching function*, denoted by $\boldsymbol{\mu}= \mathcal{G}(\boldsymbol{m}, \boldsymbol{f}; \boldsymbol{\Pi})$, is a simple reduced form way of characterizing the entire distribution of marriages, $\boldsymbol{\mu}$ as a function of exogenous factors that include a vector of available single men, $\boldsymbol{m}$, a vector of available single women, $\boldsymbol{f}$ and a matrix of parameters, $\boldsymbol{\Pi}$. $\boldsymbol{m}$ denotes an $(Z \times 1)$ vector of available men by type. The ith element of this vector, m_i denotes the stationary number of available age i men. Similarly, $\boldsymbol{f}$ denotes the $(Z \times 1)$ vector of available women where the jth element of this vector, f_j denotes the stationary number of available age j women. Let $\boldsymbol{\Pi}$ be an $(Z \times Z)$ matrix of parameters, its (i, j)th element is denoted by Π_{ij}. $\mathcal{G}()$ is a function that returns an $(Z \times Z)$ matrix of marriages $\boldsymbol{\mu}$, whose (i, j)th element, μ_{ij} gives the number of marriages between type i men and type j women.

The *marriage matching function* is given by,

$$\mu_{ij} \quad = \quad \Pi_{ij}\sqrt{m_i f_j}\prod_{k=0}^{z_{ij}} \left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k}f_{j+k}}\right)^{\frac{1}{2}(\beta S)^k}. \tag{1}$$

The term $z_{ij} = Z - \max(i, j) \geq 0$ represents the maximum length of a match before one of the spouse in the match passes away at the terminal age Z. S denotes the probability a marriage survives in any period[5] and β is the per period discount factor. μ_{i0} and μ_{0j} denote the number of type i men and type j women who chose to remain single at age i

[5]For ease of exposition, the survival rate (which is one less the divorce rate) of a marriage has been assumed to be constant across types of matches and tenure of marriage. This can easily be relaxed. While the functional form of the matching function will change, the qualitative results remain the same.

and j respectively. This equation also needs to satisfy a set of accounting constraints.[6]

I will show that the structural parameter $\ln \Pi_{ij}$ can be interpreted as the present discounted value from an (i, j) marriage relative to the present discounted per-period utility for the couple from being single forever. I outline an estimation strategy to point identify the matrix of parameters $\boldsymbol{\Pi}$ given a vector of aggregate matches and available individuals, $(\boldsymbol{\mu}, \boldsymbol{m}, \boldsymbol{f})$. These parameters are structural in that they are invariant to marriage market demand and supply changes and capture the preferences of individuals in the market. For the empirical application of this model, practitioners are often interested in the *inverse* problem.[7] That is, given an estimated vector $\boldsymbol{\Pi}$ consistent with a vector of aggregate matches $(\boldsymbol{\mu}, \boldsymbol{m}, \boldsymbol{f})$, satisfying equation (1) and the accounting constraints, how do changes in the vector of available men and women affect the distribution of matches. I will show that conditional on the exogenous vectors $(\boldsymbol{\Pi},\ \boldsymbol{m}^*,\ \boldsymbol{f}^*)$ a solution to the marriage matching function exist. Using the representation of the equilibrium in terms of multinomial probabilities, I propose a bootstrap procedure to deriving the standard errors for the estimated preferences. I also propose a simple test for the model. In the empirical application, I use this model to analyze the change in gains from marriage in the US between 1970 and 1990. This result is compared with the static matching model of Choo and Siow (2006).

[6]The accounting constraints are

$$\begin{aligned}
\mu_{0j} + \sum_{i=1}^{Z} \mu_{ij} \qquad &= f_j \quad \forall\, j, \\
\mu_{i0} + \sum_{j=1}^{Z} \mu_{ij} \qquad &= m_i \quad \forall\, i, \\
\mu_{0j}, \mu_{i0}, \mu_{ij} \quad &\geq 0 \text{ for all } i \text{ and } j.
\end{aligned}$$

I will describe these constraints in more detail in Section 2.4.

[7]Decker, Lieb, McCann and Stephens(2010) used the term *Choo-Siow Inverse Problem* to describe a similar problem in the static marriage matching model of Choo and Siow (2006).

2.2 *Assumptions*

The proposed framework employs the dynamic discrete choice framework introduced by Rust (1987). Agents are horizontally differentiated into types that change over time.[8] Specifically, I focus on type as defined by age and attempt to empirically characterize the marriage distribution by age.[9]

Stationarity: Consider a stationary society populated by overlapping generations of adults. For expositional convenience, I assume that each individual lives for Z periods irrespective of gender.[10] 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, and let m_i and f_j denote the numbers of single males of age i and females of age j at the beginning of each period. The society is stationary in the sense that the vector single men and women $\{m_i\}_{i=1}^{Z}$ and $\{f_j\}_{j=1}^{Z}$ are exogenous and taken as given.[11]

State Variables: Any **single** type i male indexed g (or type j female indexed G) in each period is characterized by two **state variables**,

- i (or j) $\in \{1, \ldots, Z\}$ is his (or her) age when single, and
- $\boldsymbol{\epsilon}_{ig}$, (or $\boldsymbol{\epsilon}_{jG}$,) is a $(Z+1)$ vector of i.i.d idiosyncratic payoffs or match specific errors specific to type i male individual, g (or type j female, G), that is unobserved to

[8] Sautmann (2011) extends the Shimer and Smith's (2000) transferable utility model of search and matching to allow for types (defined by age) that change continuously over time. She derives conditions for positive and negative assortative matching and differential age matching.

[9] While age is clearly not a complete characterization of type, the proposed methodology can be easily extended to deal with fixed horizontal attributes or types such as race, religion and education.

[10] This assumption can be relaxed to allow for differential mortality by age and gender without changing the qualitative results of the model.

[11] This is a restrictive assumption. An earlier paper, Choo and Siow (2005) take an more general equilibrium approach and allows for endogenous vectors of single men and women within a different dynamic model of marriage. This paper posits a set of accounting equations in which the number of single men and women are endogenously determined by their behavioral dynamic model.

the econometrician.

At each period a single type i male g (or single type j female G) 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. At the beginning of each period, the single male g gets to observe the $(Z+1) \times 1$ vector of idiosyncratic payoffs $\boldsymbol{\epsilon}_{ig} = \{\epsilon_{i0g}, \epsilon_{i1g}, \ldots \epsilon_{iZg}, \}$ before deciding on a utility-maximizing decision. Similarly single female G observes a $[(Z+1) \times 1]$ vector of idiosyncratic payoffs $\boldsymbol{\epsilon}_{jG}$. Single adults can choose whether or not to marry. These type specific idiosyncratic draws do not depend on the identity of the spouse the single decision maker meets or matches with.

Actions: a_{ig} (or a_{jG}) denote the action of a single type i male g (or single type j female G) where a_{ig} (or a_{jG}) $\in \{0, 1, \ldots, Z\}$. If he (or she) chooses to remain single, $a_{ig} = 0$ (or $a_{jG} = 0$), else if he (or she) chooses to match with a type k spouse, $a_{ig} = k$ (or $a_{jG} = k$).

Exogenous Parameters: The time discount factor is denoted by $\beta \in (0, 1)$. Marriages may end in divorce or the death of a spouse. Divorce occurs at some exogenous rate, δ. I assume that $\delta = 0$ in the first year of marriage for all (i, j) pairs and $\delta = 1$ in the kth year where $\max(i, j) + k > Z$. If divorce occurs in period k of a marriage, where $1 \leq k \leq Z - \max(i, j)$, the individuals g and G reenter the marriage market as single individuals of age $i + k$ and $j + k$ respectively. Let the survival probability of marriage be denoted by $S = 1 - \delta$. [12] I do not distinguish the previous marital status of the single men and women.

AS and CI: The specification of preferences over partners satisfy two assumptions: the *Additive Separability* and *Conditional Independence* assumptions. Both these assumptions were introduced by Rust (1987) in the context of a single agent dynamic discrete choice model. Let the utility of a single male g with state vector $(i, \boldsymbol{\epsilon}_{ig})$ from action a_{ig} be denoted by $v(a_{ig}, i, \boldsymbol{\epsilon}_{ig})$. The utility of a single female individual G with state

[12] The formulation carries through when I allow for duration dependence in the divorce hazard.

vector $(j, \boldsymbol{\epsilon}_{jG})$ from action a_{jG} be denoted by $w(a_{jG}, j, \boldsymbol{\epsilon}_{jG})$. I assume that the following assumptions hold:

Assumption AS *Additive Separability:*

The utility functions $v(a_{ig}, i, \boldsymbol{\epsilon}_{ig})$ and $w(a_{jG}, j, \boldsymbol{\epsilon}_{jG})$ have additively separable decompositions of the form,

$$v(a_{ig}, i, \boldsymbol{\epsilon}_{ig}) = v_a(i) + \epsilon_{iag}, \tag{2}$$

$$w(a_{jG}, j, \boldsymbol{\epsilon}_{jG}) = w_a(j) + \epsilon_{jaG}, \tag{3}$$

where ϵ_{iag} and ϵ_{jaG} are the ath component of the vector $\boldsymbol{\epsilon}_{ig}$ and $\boldsymbol{\epsilon}_{jG}$ respectively.

Assumption CI *Conditional Independence:*

The transition probability of the state variables for males and females respectively factorize as

$$\mathbb{P}\{i', \boldsymbol{\epsilon}'_{ig} \mid i, \boldsymbol{\epsilon}, a\} = h(\boldsymbol{\epsilon} \mid i) \cdot \mathcal{F}_a(i' \mid i), \tag{4}$$

$$\mathbb{P}\{j', \boldsymbol{\epsilon}'_{jG} \mid j, \boldsymbol{\epsilon}, a\} = h(\boldsymbol{\epsilon} \mid i) \cdot \mathcal{R}_a(j' \mid j), \tag{5}$$

where $h(\boldsymbol{\epsilon})$ is the multivariate pdf of the i.i.d $\boldsymbol{\epsilon}$, and $\mathcal{F}_a(i' \mid i)$ (or $\mathcal{R}_a(j' \mid j)$) is the probability that the male (or female) individual will next be single again at i' (or j') given action a and his (or her) current age i (or j).

I also maintain the assumption introduced in Choo and Siow (2006) that $\boldsymbol{\epsilon}_{ig}$ is drawn from MacFadden's type I extreme value distribution.[13]

The **CI** limits the dependence structure on the state variables. As discussed in Rust (1994), it says that the observed states i' (and j') are sufficient statistics for the unobserved states $\boldsymbol{\epsilon}'_{ig}$ (and $\boldsymbol{\epsilon}'_{jG}$). Any dependence between $\boldsymbol{\epsilon}'$ and $\boldsymbol{\epsilon}$ is transmitted through ages i' and j'. $\mathcal{F}_a(i' \mid i)$ is the transition probability that a type i male g will next find himself single at age i' given his action a at age i. Similarly $\mathcal{R}_a(j' \mid j)$ is the transition probability that a type j female G will next find herself single at age j' given her action

[13] The marginal density is given by $h(\boldsymbol{\epsilon}_{ig} \mid i) = \prod_{a_{ig}=0}^{Z} \exp[-\epsilon_{iag} + c] \exp[-\exp(-\epsilon_{iag} + c)]$, where c is the Euler constant.

a. $\boldsymbol{\epsilon}$ are iid noise that are superimposed on this process. I will go into the details of the structure of the utilities in the next section.

$\mathcal{F}_a(r \mid i)$ is the transition probability that a type i male g will next find himself single at age r given his action a. Clearly $\mathcal{F}_a(r \mid i) = 0$ for all $r \leq i$ and all a. If g chooses to be single $a = 0$, $\mathcal{F}_0(r \mid i) = 1$ for $r = i + 1$ and zero elsewhere. If g chooses to match with a type j spouse (a=j), $\mathcal{F}_a(r \mid i)$ takes the form, For $i < r \leq Z$,

$$\mathcal{F}_a(r \mid i) = \begin{cases} \delta(\beta S)^{r-(i+1)}, & \text{if} \quad i+1 \leq r \leq i + z_{ij} \\ (\beta S)^{r-(i+1)}, & \text{if} \quad r > i + z_{ij}. \end{cases} \tag{6}$$

Similarly $\mathcal{R}_a(r \mid j)$ is the transition probability that a type j female G will next find herself single at age r given her action a. $\mathcal{R}_a(r \mid j) = 0$ for all $r \leq j$ and all a. For action $a = 0$, $\mathcal{R}_0(r \mid j) = 1$ for $r = j + 1$ and zero elsewhere. If she chooses to match with a type i spouse (a=i), $\mathcal{R}_a(r \mid j)$ takes the form,

$$\mathcal{R}_a(r \mid j) = \begin{cases} \delta(\beta S)^{r-(j+1)}, & \text{if} \quad j+1 \leq r \leq j + z_{ij} \\ (\beta S)^{r-(j+1)}, & \text{if} \quad r > j + z_{ij}. \end{cases} \tag{7}$$

Functional Form of Utilities: The model adopts a full commitment framework; the decision to marry locks an individual into a stream of payoffs in the event that the marriage does not dissolve due to divorce or death of either spouse. Let α_{ijk} be the $k'th$ period marital output accrued to a type i male when married to a type j female today. Similarly γ_{ijk} be the $k'th$ period marital output accrued to a type j female when married to a type i male.

Suppose male g (or female G) chooses to marry an age j female (or i male), his (or her) one period utility functions (given by Equations (2) and (3) above) respectively are

$$v(a_{ig} = j, i, \boldsymbol{\epsilon}_{ig}) \;=\; \begin{cases} \boldsymbol{\alpha}_i(j) - \tau_{ij} + \epsilon_{ijg}, & \text{if} \quad 1 \leq a \leq Z \\ \alpha_{i0} + \epsilon_{i0g}, & \text{if} \quad a = 0, \quad \text{and} \end{cases} \tag{8}$$

$$w(a_{jG} = i, j, \boldsymbol{\epsilon}_{jG}) \;=\; \begin{cases} \boldsymbol{\gamma}_j(i) + \tau_{ij} + \epsilon_{ijG}, & \text{if} \quad 1 \leq a \leq Z \\ \gamma_{0j} + \epsilon_{0jG}, & \text{if} \quad a = 0, \end{cases} \tag{9}$$

where the present discounted gains from the match $\boldsymbol{\alpha}_i(j)$ and $\boldsymbol{\gamma}_j(i)$ take the form,

$$\boldsymbol{\alpha}_i(j) = \sum_{k=0}^{z_{ij}} (\beta S)^k \, \alpha_{ijk}, \text{ and } \boldsymbol{\gamma}_j(i) = \sum_{k=0}^{z_{ij}} (\beta S)^k \, \gamma_{ijk}. \tag{10}$$

α_{i0} and γ_{0j} are the per-period utilities from remaining single for i type males and j type females respectively. Recall that $z_{ij} = Z - \max(i, j)$ captures the maximum length of the match given the terminal age of Z.

Equation (8) says that if g marries a type j women, he receives the mean utility from the match equal to $\boldsymbol{\alpha}_i(j) - \tau_{ij}$, plus an idiosyncratic shock ϵ_{ijg}. The mean utility depends only on the type of men and women in the match and does not depend on the precise identity of the spouse or the decision maker g. $\sum_{k=0}^{z_{ij}} (\beta S)^k \, \alpha_{ijk}$ captures the present discounted stream of male marital payoffs in the event that the marriage does not dissolve. In choosing this match, g commits to pay a once off transfer, τ_{ij} specific to these two types of individuals matching. Similarly in equation (9), single females G of type j who decide to marry type i men agree to receive this equilibrium transfer. In accepting the match, she locks herself to a stream of marital payoffs, of which the present discounted value equals $\sum_{k=0}^{z_{ij}} (\beta S)^k \, \gamma_{ijk}$. So, if individual g of type i wants to marry a woman G of type j, he has to transfer τ_{ij} 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} 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. This transfer can be positive or negative. 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.

2.3 *The Agents' Decision Problem*

The single male at age i chooses a sequence of decisions $\boldsymbol{a}_{ig} = \{a_{ig}, a_{i+1g}, \dots a_{Zg}\}$, where $a_{ig} = a(i, \boldsymbol{\epsilon}_{ig}) \in \{0, 1, \dots, Z\}$ is the expected discounted utilities maximizing choice g

makes in the event that he is single at age i. The value function $V_\alpha(i, \boldsymbol{\epsilon}_g)$ is defined by

$$V_\alpha(i, \boldsymbol{\epsilon}_{ig}) = \max_{\boldsymbol{a}} \mathbb{E}\Big\{ \sum_{k=i}^{Z} \beta^{k-i}\Big([\boldsymbol{\alpha}_k(a_{kg}) - \tau_{ka}]\mathbb{I}(a_{kg} \neq 0) + \alpha_{k0}\mathbb{I}(a_{kg} = 0) + \epsilon_{kag}\Big) \mid i, \boldsymbol{\epsilon}_{ig}\Big\}.$$

Similarly for single age j women G, her value function takes the form

$$W_\gamma(j, \boldsymbol{\epsilon}_{jG}) = \max_{\boldsymbol{a}'} \mathbb{E}\Big\{ \sum_{k=j}^{Z} \beta^{k-i}\Big([\boldsymbol{\gamma}_k(a'_{kG}) + \tau_{ak}]\mathbb{I}(a'_{kG} \neq 0) + \gamma_{0k}\mathbb{I}(a'_{kG} = 0) + \epsilon_{akG}\Big) \mid j, \boldsymbol{\epsilon}_{jG}\Big\}.$$

where $\boldsymbol{a}'_{jG} = \{a'_{jG}, a'_{j+1G}, \ldots, a'_{ZG}\}$. The Bellman equations for male g and female G respectively takes the familiar form

$$\begin{aligned} V_\alpha(i, \boldsymbol{\epsilon}_g) = \quad \max{}_a \quad & \Big\{[\boldsymbol{\alpha}_i(a) - \tau_{ia}]\mathbb{I}(a \neq 0) + \alpha_{i0}\mathbb{I}(a = 0) + \epsilon_{iag} \\ & + \beta\mathbb{E}\big(V_\alpha(i', \boldsymbol{\epsilon}'_g) \mid i, \boldsymbol{\epsilon}_g, a\big)\Big\}, \end{aligned}$$

$$\begin{aligned} W_\gamma(j, \boldsymbol{\epsilon}_G) = \quad \max{}_{a'} \quad & \Big\{[\boldsymbol{\gamma}_j(a') + \tau_{a'j}]\mathbb{I}(a' \neq 0) + \gamma_{0j}\mathbb{I}(a' = 0) + \epsilon_{a'jG} \\ & + \beta\mathbb{E}\big(W_\gamma(j', \boldsymbol{\epsilon}'_G) \mid j, \boldsymbol{\epsilon}_G, a'\big)\Big\}. \end{aligned}$$

Integrated Value Functions: Rust (1987) showed that Assumptions AS and CI allow the Bellman equations be represent in a form where the unobservables are integrated out. Let $\boldsymbol{V}_i$ and $\boldsymbol{W}_j$ be the corresponding integrated value function for a single age i male and j female respectively. That is $\boldsymbol{V}_i = \mathbb{E}V_{\boldsymbol{\alpha}}(i, \boldsymbol{\epsilon}_g) = \int V_{\boldsymbol{\alpha}}(i, \boldsymbol{\epsilon}_g)\, dH(\boldsymbol{\epsilon}_g)$ and $\boldsymbol{W}_j = \mathbb{E}W_{\boldsymbol{\gamma}}(j, \boldsymbol{\epsilon}_G) = \int W_{\boldsymbol{\gamma}}(j, \boldsymbol{\epsilon}_G)\, dH(\boldsymbol{\epsilon}_G)$. The integrated Bellman equation for a single type i male and type j female then takes the form,

$$\begin{aligned} \boldsymbol{V}_i = \int \quad \max{}_{a \in \mathcal{D}} \quad & \Big\{[\boldsymbol{\alpha}_i(a) - \tau_{ia}]\mathbb{I}(a \neq 0) + \alpha_{i0}\mathbb{I}(a = 0) + \epsilon_{iag} \\ & + \beta \sum_{i'} \mathcal{F}_a(i' \mid i) \cdot \boldsymbol{V}_{i'}\Big\} h(d\epsilon), \end{aligned} \tag{11}$$

$$\begin{aligned} \boldsymbol{W}_j = \int \quad \max{}_{a \in \mathcal{D}} \quad & \Big\{[\boldsymbol{\gamma}_j(a) + \tau_{aj}]\mathbb{I}(a \neq 0) + \gamma_{0j}\mathbb{I}(a = 0) + \epsilon_{ajG} \\ & + \beta \sum_{j'} \mathcal{R}_a(j' \mid j) \cdot \boldsymbol{W}_{j'}\Big\} h(d\epsilon). \end{aligned} \tag{12}$$

Consider decomposing the integrated value function $\boldsymbol{V}_i$ and $\boldsymbol{W}_j$ in equations (11) and (12) into a mean component that is dependent on the utility maximizing choice and an idiosyncratic component. This decomposition together with the distributional assumption provides a closed form representation for the conditional choice probabilities of a particular type of spouse.

The mean component also referred to as the *choice specific value functions* for type i males and j females is denoted by $\tilde{v}_{ij}$ and $\tilde{w}_{ij}$ respectively. They are

$$\begin{aligned}\tilde{w}_{ij} &= [\boldsymbol{\gamma}_j(i)+\tau_{ij}]\mathbb{I}(i\neq 0)+\gamma_{0j}\mathbb{I}(i=0)+\sum_{j'}\mathcal{R}_i(j'\mid j)\cdot\boldsymbol{W}_{j'}\\ \tilde{v}_{ij} &= [\boldsymbol{\alpha}_i(j)-\tau_{ij}]\mathbb{I}(j\neq 0)+\alpha_{i0}\mathbb{I}(i=0)+\sum_{i'}\mathcal{F}_j(i'\mid i)\cdot\boldsymbol{V}_{i'}.\end{aligned}$$

This provides an alternative representation for male g's and female G's optimization problem that is convenient when talking about stability in Section 2.7.[14] The value functions can now be written as

$$V(i,\boldsymbol{\epsilon}_{ig},\boldsymbol{\alpha}) = \max_{a\in\mathcal{D}}\ \left\{\tilde{v}_{ia}+\epsilon_{iag}\right\} \tag{13}$$

$$W(j,\boldsymbol{\epsilon}_{jG},\boldsymbol{\gamma}) = \max_{a\in\mathcal{D}}\ \left\{\tilde{w}_{aj}+\epsilon_{ajG}\right\} \tag{14}$$

Let the conditional choice probability $\mathcal{P}_{ij}$ denote the probability that choice j is the optimal choice for males at age i, that is $\mathcal{P}_{ij}=\int\mathbb{I}\{j=\arg\max_{a\in\mathcal{D}}(\tilde{v}_{ia}+\epsilon_{iag})\}h(d\boldsymbol{\epsilon})$. Similarly for females, $\mathcal{Q}_{ij}$ is the probability that choice i is the optimal choice for females at age j. That is, $\mathcal{Q}_{ij}=\int\mathbb{I}\{i=\arg\max_{a\in\mathcal{D}}(\tilde{w}_{aj}+\epsilon_{ajG})\}h(d\boldsymbol{\epsilon})$. The conditional choice probability can be expressed as a function of the normalized choice value functions, $(\tilde{v}_{ij}-\tilde{v}_{i0})$. In our case, the probability that a type i male who matches with a type j female will have the familiar multinomial logit form,

$$\mathcal{P}_{ij} = \frac{\exp(\tilde{v}_{ij}-\tilde{v}_{i0})}{1+\sum_{r=1}^{Z}\exp(\tilde{v}_{ir}-\tilde{v}_{i0})}, \tag{15}$$

and similarly for females,

$$\mathcal{Q}_{ij} = \frac{\exp(\tilde{w}_{ij}-\tilde{w}_{0j})}{1+\sum_{r=1}^{Z}\exp(\tilde{w}_{rj}-\tilde{w}_{0j})}. \tag{16}$$

[14] Browning, Chiappori and Weiss (2007) and Chiappori, Salanié and Weiss (2010) provided a proof of the existence of $\tilde{v}_{ij}$ and $\tilde{w}_{ij}$. Berry (1994) also provided a proof of uniqueness in the context of differentiated goods demand models.

In this finite horizon case, the integrated value functions for male and females also take a convenient recursive structure,

$$\boldsymbol{V}_i = \begin{cases} \alpha_{i0} + c + \beta \boldsymbol{V}_{i+1} - \ln \mathcal{P}_{i0} & : \quad i < Z \\ \alpha_{i0} + c - \ln \mathcal{P}_{i0} & : \quad i = Z \end{cases} \tag{17}$$

$$\boldsymbol{W}_j = \begin{cases} \gamma_{0j} + c + \beta \boldsymbol{W}_{j+1} - \ln \mathcal{Q}_{0j} & : \quad j < Z \\ \gamma_{0j} + c - \ln \mathcal{Q}_{0j} & : \quad i = Z \end{cases} \tag{18}$$

Equation 17 says that the expected value of participating in the marriage market at age i, $\boldsymbol{V}_i$ can be divided into two components. The first is the expected utility from being single this period which is comprised of $\alpha_{i0} + c$ and the expected value of participating in the marriage market next period as an older $(i+1)$ individual 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 locked into marriage is high.

2.4 *Equilibrium and the Dynamic Marriage Matching Function*

The log-odds ratio of an (i,j) match relative to i remaining single in can also be expressed in terms of normalized choice specific value functions, $\tilde{v}_{ij} - \tilde{v}_{i0}$. It describes the expected payoffs for an i type male marrying a j type female relative to remaining single that period. This is given by

$$\log \left\{ \frac{\mathcal{P}_{ij}}{\mathcal{P}_{i0}} \right\} = \boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_i(0) - \tau_{ij} - \sum_{k=1}^{z_{ij}} (\beta S)^k \left(c + \ln \mathcal{P}_{i+k,0}^{-1} \right) \tag{19}$$

The term $\boldsymbol{\alpha}_i(j) = \sum_{k=1}^{z_{ij}} (\beta S)^k \alpha_{ijk}$ represents the stream of expected period utility he gets from the match in the event that the marriage does not dissolve. Abusing notation slightly, $\boldsymbol{\alpha}_i(0)$ denotes $\sum_{k=1}^{z_{ij}} (\beta S)^k \alpha_{i+k,0}$. In the event of divorce or death of a spouse at age $(i+k) < Z$ where $0 < k \leq z_{ij}$, his expected value of being single is $\alpha_{i+k} + c + \ln \mathcal{P}_{i+k,0}^{-1} + \beta \boldsymbol{V}_{i+k+1}$. By repeated use of the recursive equations (17) and (18), it can be shown that at each age $i+k$ of marriage, the difference in expected utility from being locked in marriage and participating in the marriage market that period is represented by $\alpha_{i_k,j+k} - \alpha_{i+k,0} - \beta(1-\delta)(c + \ln \mathcal{P}_{i+k,0}^{-1})$. $\ln \mathcal{P}_{i+k,0}$ is a statistic for the gains from participating in the marriage market at $i+k$. In choosing this match, the i type male commits to pay a one-time match specific transfer τ_{ij} to his spouse.

Similarly for females, the log-odds ratio that a j type female marries an i type male relative to remaining single equals the difference in choice specific value functions $\tilde{w}_{ij} - \tilde{w}_{i0}$. That is,

$$\log\left\{\frac{\mathcal{Q}_{ij}}{\mathcal{Q}_{0j}}\right\} = \boldsymbol{\gamma}_j(i) - \boldsymbol{\gamma}_j(0) + \tau_{ij} - \sum_{k=1}^{z_{ij}} (\beta S)^k \left(c + \ln \mathcal{Q}_{0,j+k}^{-1}\right) \tag{20}$$

Equation (20) gives the difference in systematic expected payoffs for a j type female marrying an i type male relative to remaining single that period. The interpretation of the various terms in (20) is analogous to that for males. Recall that γ_{ij}^k is the females share of 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. When the marriage does not dissolve, the female share of the discounted within marriage payoffs is $\boldsymbol{\gamma}_j(i) = \sum_{k=1}^{z_{ij}} (\beta S)^k \gamma_{ijk}$ If she agrees to this match, she receives an equilibrium transfer τ_{ij} from her partner. Given that divorce occurs at an exogenous rate δ, the term $\delta(\gamma_{0,j+k} + c + \beta \boldsymbol{W}_{0,j+k+1} + \ln \mathcal{Q}_{0,j+k}^{-1})$ captures the expected value of re-entering the marriage market in the future in the event of divorce. 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 (20).

Rearranging the terms of the the log-odds ratios in Equations (19) and (20) delivers a system of $(Z \times Z)$ quasi-demand and quasi-supply equations respectively.

$$\ln \mathcal{P}_{ij} - \sum_{k=0}^{z_{ij}} (\beta S)^k \ln \mathcal{P}_{i+k,0} = \boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_i(0) - \tau_{ij} - \kappa \tag{21}$$

$$\ln \mathcal{Q}_{ij} - \sum_{k=0}^{z_{ij}} (\beta S)^k \ln \mathcal{Q}_{0,j+k} = \boldsymbol{\gamma}_j(i) - \boldsymbol{\gamma}_j(0) + \tau_{ij} - \kappa. \tag{22}$$

The constant κ is the geometric sum of Euler's constants, $\kappa = c\beta S(1 - (\beta S)^{z_{ij}})/(1 - \beta S)$.[15] The left hand side of equation (21) is the log of an (i,j) choice probabilities

[15]In the static framework of Choo and Siow (2006), we get an analogous representation of quasi-

$\mathcal{P}_{ij}$, scaled by the weighted average of the probabilities of remaining single in the future $\sum_{k=0}^{z_{ij}}(\beta S)^k \ln \mathcal{P}_{i+k,0}$. The denominator represents the opportunity cost of future participation in the marriage market that an i type male incurs when he chooses to match with a j type female. When the probability of remaining single in the future, $\mathcal{P}_{i+k,0}$ is large, then the forgone opportunity of being locked into marriage is small and vice versa. This ratio equals the difference in present discounted utilities from being locked into an (i, j) relative to the present discounted value of being single forever, $(\boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_i(0))$ less the equilibrium transfer, τ_{ij} that needs to be paid. The interpretation for equation (22) in terms of female match probabilities is similar. The central difference being that the type j female is the recipient of the transfer.

Definition 1: A *marriage market equilibrium* consists of a vector of males, $\boldsymbol{m}$ and females, $\boldsymbol{f}$ across individual type, the vector of marriage $\boldsymbol{\mu}$, and the vector of transfers, $\boldsymbol{\tau}$ 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 combinations of (i, j). That is, for each of the $(Z \times Z)$ sub-markets,

$$m_i \mathcal{P}_{ij} = f_j \mathcal{Q}_{ij} = \mu_{ij} \quad \blacksquare$$

Dynamic Marriage Matching Function: Let p_{ij} and q_{ij} denote the maximum likelihood estimators of the probability that type i male matches with type j female, $\mathcal{P}_{ij}$ and type j female matches with i male, $\mathcal{Q}_{ij}$ respectively. That is, $p_{ij} = \mu_{ij}/m_i$ and $q_{ij} = \mu_{ij}/f_j$. The above marriage market clearing conditions and the ML estimators for the choice probabilities is applied to the system of quasi-supply and demand equations, (21) and (22) respectively to derive the Dynamic Marriage Matching Function for an

demand and quasi-supply of spouses corresponding to the case when $z_{ij} = 0$. That is

$$\begin{aligned}
\ln \mathcal{P}_{ij} - \ln \mathcal{P}_{i,0} &= \alpha_{ij} - \alpha_{i,0} - \tau_{ij} \\
\ln \mathcal{Q}_{ij} - \ln \mathcal{Q}_{0,j} &= \gamma_{ij} - \gamma_{0,j} + \tau_{ij}
\end{aligned}$$

(i,j) marriage (when $z_{ij} > 0$)[16] given by Equation (1) shown earlier,

$$\mu_{ij} \quad = \quad \Pi_{ij}\sqrt{m_i f_j}\prod_{k=0}^{z_{ij}}\left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k}f_{j+k}}\right)^{\frac{1}{2}(\beta S)^k}$$

where $\ln \Pi_{ij} = \frac{1}{2}\big(\boldsymbol{\alpha}_i(j) + \boldsymbol{\gamma}_j(i) - \boldsymbol{\alpha}_i(0) - \boldsymbol{\gamma}_j(0)\big) - \kappa$.

The dynamic marriage matching function also needs to satisfy the accounting constraints given by Equations (23), (24) and (25):

$$\mu_{0j} + \sum_{i=1}^{Z}\mu_{ij} = f_j \;\forall\; j \tag{23}$$

$$\mu_{i0} + \sum_{j=1}^{Z}\mu_{ij} = m_i \;\forall\; i \tag{24}$$

$$\mu_{0j}, \mu_{i0}, \mu_{ij} \geq 0 \;\forall\; i,j \tag{25}$$

Equation (23) says that the total number of j type women who marry and the number of unmarried j type women must be equal to the number of available j type women for all j. Similarly Equation (24) says that the total number of women who marry i type men and the number of unmarried i type men must be equal to the number of available i type men for all i. Equation (25) holds because the number of unmarrieds of any type and gender, and the number of marriages between type i men and type j women must be non-negative.

Given the preference parameters of the system, $\boldsymbol{\Pi}$, practitioners are often interested in how variations in the supply population vectors, $\boldsymbol{m}$ and $\boldsymbol{f}$, affect the distribution of marriages as represented by $\boldsymbol{\mu}$. I'll refer to this as the DMM (Dynamic Marriage Matching) Inverse Problem. A formal statement of this problem follows:

Definition 2: - *Dynamic Marriage Matching (DMM) Inverse Problem*
Given a matrix of preferences $\boldsymbol{\Pi}$, whose elements are non-negative and strictly positive

[16]This condition ensures that neither spouse is at a terminal age. If $z_{ij} = 0$, the Dynamic Marriage Matching Function reduces to the static marriage matching function of Choo and Siow (2006), that is $\mu_{ij} = \Pi_{ij}\sqrt{\mu_{i0}\mu_{0j}}$.

population vectors, $\boldsymbol{m}$ and $\boldsymbol{f}$, does there exist a unique non-negative marital distribution $\boldsymbol{\mu}$ that is consistent with $\boldsymbol{\Pi}$, that satisfies equations (23), (24), (25) and (20). ∎

Taking Π_{ij}, $\boldsymbol{m}$ and $\boldsymbol{f}$ as exogenously given, Equation (1) defines a $I \times J$ system of polynomials with the $I \times J$ elements of μ as unknowns. Like in Choo-Siow (2006), the model can be reformulated to an $I + J$ system with $I + J$ number of unmarrieds of each type, μ_{i0} and μ_{0j}, as unknowns. This reduced system defined by equations (26) and (27) below is derived by summing Equation (1) over all i's and Equation (1) over all j's respectively. After solving for μ_{i0} and μ_{0j}, then the μ_{ij}'s are fully determined by Equation (1).

$$m_i - \mu_{i0} = \sum_{i=1}^{I} \Pi_{ij} \sqrt{m_i f_j} \prod_{k=0}^{z_{ij}} \left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k} f_{j+k}} \right)^{\frac{1}{2}(\beta S)^k} \tag{26}$$

$$f_j - \mu_{0j} = \sum_{j=1}^{J} \Pi_{ij} \sqrt{m_i f_j} \prod_{k=0}^{z_{ij}} \left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k} f_{j+k}} \right)^{\frac{1}{2}(\beta S)^k}. \tag{27}$$

2.5 *Existence and Uniqueness*

As noted in Chiappori, McCann and Nesheim (2009), transferable utility marriage matching model is equivalent to an optimal transportation (Monge-Kantorovich) linear programming problem.[17] They showed that optimal assignment in (Monge-Kantorovich) linear programming problem corresponds to stable matching and that optimal assignment are shown to exist under mild conditions. This equivalence brings to bear the wide body of knowledge about linear programming and optimal transportation. Despite the complication arising from the dynamics, the formulation of the marriage matching model in this paper reduces to a structure identical to that introduced by Choo and

[17] A continuum version of the Sharpley and Shubik's(1972) transferable utility assignment model has also been analyzed by Gretszky, Ostroy and Zame (1992, 1999). They proved that the transferable utility assignment model can either be modeled as a linear programming problem, a coorperative game or an exchange economy. They proved the equivalence of these three views in a setting with a continuum of agents. Agents in this economy are endowed with a single contract and are not free to trade any contracts. This restriction is relaxed by Chiappori, McCann and Nesheim (2009).

Siow (2006) and extended in Chiappori, Salaniè and Weiss (2011). The general existence result holds in this model. Readers should refer to Chiappori, McCann and Nesheim (2009) for precise statements.

As for uniqueness, linear programming models on compact convex feasible set generically have unique solutions. However for finite population, stable matching is generally not unique. It is possible to marginally alter the individual payoffs without violating the conditions for stability. In the limit as the population becomes large, uniqueness is established. I again refer interested readers to Chiappori, McCann and Nesheim (2009) for more precise statements.

2.6 *Identification*

2.6.1 Dynamic Gains to Marriage

Choo and Siow (2006) introduced a statistic for the *Total Gains* to an (i, j) marriage relative to remaining single. It is the ratio of the number of (i, j) to the geometric mean of the number of unmarrieds of each type, that is

$$\ln \frac{\mu_{ij}^2}{\mu_{i0}\mu_{0j}} = \ln \frac{p_{ij}q_{ij}}{p_{i0}q_{0j}} = (\alpha_{ij} + \gamma_{ij}) - (\alpha_{i0} + \gamma_{0j}) = 2 \ln \pi_{ij}. \tag{28}$$

The dynamic analogue of this statistic derived from Equation (1) takes the following form,

$$\ln \left(p_{ij}q_{ij} \Big/ \prod_{k=0}^{z_{ij}} (p_{i+k,0}q_{0j+k})^{(\beta S)^k} \right) = \boldsymbol{\alpha}_i(j) + \boldsymbol{\gamma}_j(i) - \boldsymbol{\alpha}_i(0) - \boldsymbol{\gamma}_j(0) - 2\kappa = 2 \ln \Pi_{ij}. \tag{29}$$

The interpretation of the statistic is similar to the static case. It gives the present discounted utility from being locked in an (i, j) match relative to the present discounted sum of the per period payoff from being single forever.[18] This statistic is point identified. The right hand side of (29) comprise of only the primitives of the model and are invariant to changes in the vectors of unmarried men, $\boldsymbol{m}$, and women, $\boldsymbol{f}$. It becomes the basis of the empirical application in Section 3. I will refer to the statistic $2 \ln \pi_{ij}$ as defined by Equation (28) as *Static Gains* and $2 \ln \Pi_{ij}$ from Equation (29) as *Dynamic Gains*.

[18] In other words, since the couple is committing to being single forever, the right hand side term does not include the expected opportunity cost of participating in the marriage market from being single.

2.6.2 Bootstrap Standard Errors

The representation of the matching function in Equation (29) provides a natural way of generating bootstrap standard errors for the marriage gains statistics. Given an observed distribution of $\boldsymbol{m}$, $\boldsymbol{f}$ and $\boldsymbol{\mu}$, there is an implied vector of maximum likelihood estimators $\boldsymbol{p}$ and $\boldsymbol{q}$.

To derive the standard error of the gains statistic, I first sample with replacement in blocks from the pool of individuals in each age group keeping track of whether the individual is married and to whom he or she is married to. For example, suppose there are m_1 type 1 males.[19] For each bootstrap sample s of size m_1, I generate the implied probability vector p_1^s. This is carried out for each block of i and j generating the implied vector of $\boldsymbol{p}^s$ and $\boldsymbol{q}^s$ and the matrix of statistics $\boldsymbol{\Pi}^s$. The standard error and confidence interval is computed from the demeaned sampling distribution of $\boldsymbol{\Pi}^s$.

The asymptotic distribution for the gains statistic is complicated by the covariance structure of $\mathcal{P}_{ij}$ and $\mathcal{P}_{kj}$ where $i \neq k$ which depends on the underlying assumption of the model. Choo (2012) compares the coverage of the proposed block bootstrap method with the asymptotic distribution of the dynamic gains statistic.

2.6.3 A Test of the Model

Equations (21) and (22) can be expressed in terms of the maximum likelihood estimators p_{ij} and q_{ij}. That is,

$$\ln\left(p_{ij}\Big/\prod_{k=0}^{z_{ij}} p_{i+k,0}^{(\beta S)^k}\right) = \boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_i(0) - \tau_{ij} - \kappa, \tag{30}$$

$$\ln\left(q_{ij}\Big/\prod_{k=0}^{z_{ij}} q_{0,j+k}^{(\beta S)^k}\right) = \boldsymbol{\gamma}_j(i) - \boldsymbol{\gamma}_j(0) + \tau_{ij} - \kappa. \tag{31}$$

Let

$$\begin{aligned} n_{ij}(\boldsymbol{\mu},\boldsymbol{m},\boldsymbol{f}) &= \ln\left(p_{ij}\Big/\prod_{k=0}^{z_{ij}} p_{i+k,0}^{(\beta S)^k}\right) \quad \text{and} \\ \mathcal{N}_{ij}(\boldsymbol{\mu},\boldsymbol{m},\boldsymbol{f}) &= \ln\left(q_{ij}\Big/\prod_{k=0}^{z_{ij}} q_{0,j+k}^{(\beta S)^k}\right) \end{aligned}$$

[19] That is, $\sum_{j=0}^{Z} \mu_{1j} = m_1$.

Proposition 1 below provides a simple test for our model:

Proposition 1 *Holding α_{ijk}, γ_{ijk}, and δ_{ijk} fixed for all (i,j,k), any changes in available men m_i or women f_j that leads to an increase in $n_{ij}(\boldsymbol{\mu},\boldsymbol{m},\boldsymbol{f})$ would also lead to a decrease in $\mathcal{N}_{ij}(\boldsymbol{\mu},\boldsymbol{m},\boldsymbol{f})$ and vice versa.*

In other words, any changes in relative scarcity of men and women that changes the market clearing division of surplus τ_{ij} would make n_{ij} and $\mathcal{N}_{ij}$ move in opposite directions. If our model is true, a simple regression of estimates of $\hat{n}_{ij}(\boldsymbol{\mu},\boldsymbol{m},\boldsymbol{f})$ against $\hat{\mathcal{N}}_{ij}(\boldsymbol{\mu},\boldsymbol{m},\boldsymbol{f})$ should give a slope coefficient of -1.

2.7 *Stability*

Browning, Chiappori and Weiss (2007) and Chiappori, Salaniè and Weiss (2010) introduced a stable matching characterization for preference utility that maintains the additive separability structure introduced in Choo and Siow (2006). These stable matching characterizations consist of a set of inequalities that translate naturally when defining equilibrium probabilities of different types of matches as in Choo and Siow (2006). These characterizations have significant implications in the empirical implementation of models with this additively separable structure. The dynamic marriage matching outlined here maintains the additively separable structure. Equations (13) and (14) show that despite this more complicated dynamic setting, the choice specific utilities still maintain the additively separable structure. The choice specific utilities can be decomposed into a mean choice specific value function and an i.i.d. idiosyncratic component. For completeness, the stable matching characterization of Chiappori, Salaniè and Weiss (2010) is reproduced in Lemma 2 of the appendix.

3 Empirical Application

3.1 *Changes in the Marriage distribution in the 1970s, 80s and 1990s*

The model is used to analyze the changes in the marriage distribution in the US over two decades from 1970 to 1990. From a demographic viewpoint, this period saw significant changes in the number of single men and women. In particular, the baby boomers entered a marriageable age in the 1980s and 1990s. This was also a period of major socio-political changes that affected marriage as an institution. Many have argued that federal legislative changes like the legalization of abortion and no-fault divorce have changed the gains to marriage. I construct the distribution of new marriages and single men and women for individuals aged between 16 and 75 over the two decades for the US. To minimize sparseness in the marriage distribution, a two-year distribution of new marriages is constructed (instead of a one-year). The marriage distributions by age, $\widehat{\mu}_{ij}$ for 1971/72, 1981/82 and 1991/92, is constructed using data the Vital Statistics taken from the NBER collection of the National Center for Health Statistics.[20]. These files contain a sample of the new marriage records from reporting states across the three periods.[21] The number of single men and women by age, $\widehat{\mu}_{i0}$ and $\widehat{\mu}_{0j}$, come from the 1970, 1980 and 1990 US Census. To be consistent on the data of new marriages from the Vital Statistics, only unmarried individuals from matching reporting states are included.[22] The individuals are aged between 16 and 75. An individual is considered *unmarried* is his or her marital status is not equal to (i) married spouse present, or (ii) married spouse absent. Census weights are used to get an estimate of the total unmarried counts.

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

[21]Weights in the Vital Statistics files are used to get an estimate of the total number of each type of marriage in reporting states.

[22]The reporting states are 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.

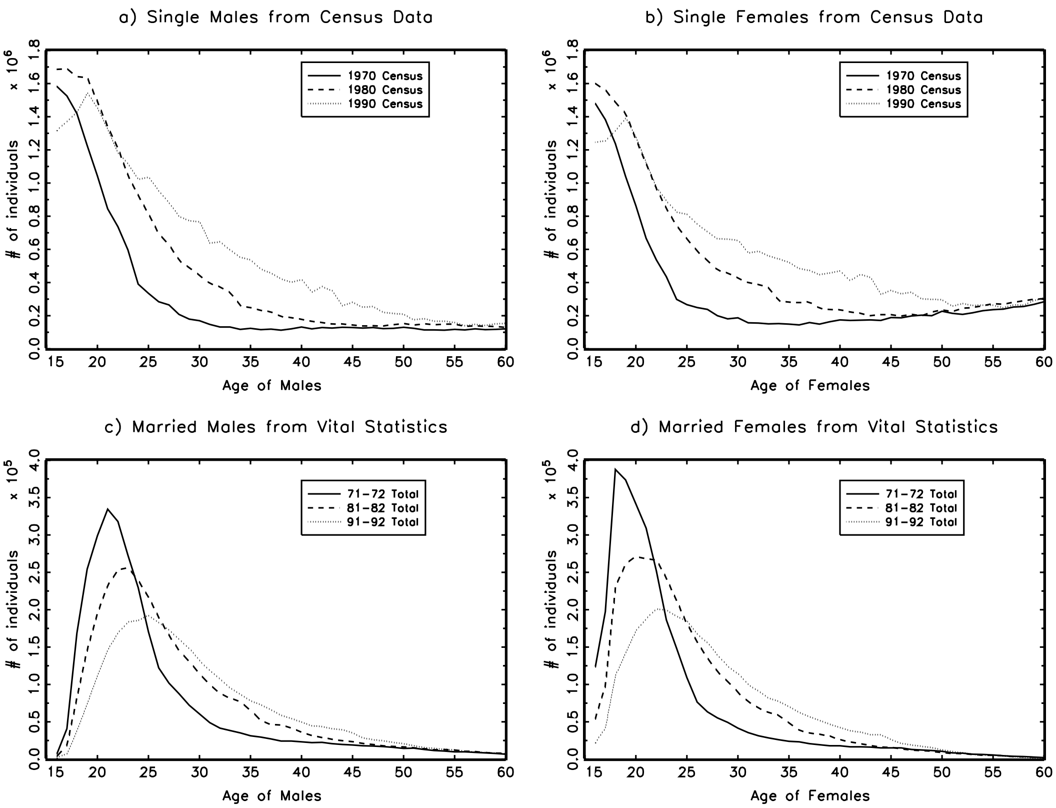

FIGURE 1

Figures 1 a) and 1 b) plot the distribution of single men and women according the the three US Census. It shows some familiar patterns. There are fewer single individuals at older ages than at younger ages. As we move from 1970 to 1980 and 1990, we see a dramatic change in the number of available single men and women as the baby boomers enter marriageable age. The gender differences in later ages arise from higher mortality rates among older men and lower remarriage rates among divorced women. Figures 1 c) and d) graph the marginal distribution of two-year new marriages over the period. There is a clear right shift in the distribution of new marriages across genders arising from more delayed marriage. The modal age of marriage across gender has also increased. For males, the modal age of marriage went from around 21 in 1971/72 to 23 in 1981/82 and 25 in 1991/92. The modal age for females which is slightly younger compared to males also increased from around 18 in 1971/72 to 20 in 1981/82 and 22 in 1991/92.

Summary statistics of the data are given in Table 1 below. According to the 1970 Census, there were 16.0 and 19.6 million single men and women respectively between the age of 16 and 75. By 1980, the number of available men and women had increased by 46.2% and 39% respectively to 23.4 and 27.2 million men and women respectively. The change in population between 1980 and 1990 was more modest. In 1990, there were 28.4 million men and 31.6 million women, an increase of 21.4% and 15.9% respectively from 1980. In the data sample constructed from the Vital Statistics, there are 3.24 million new marriages recorded in the two-years 1971-72, while in 1981-82 there are 3.45 million new marriages. This is an increase of 6.5 % compared to the around 40 % increase in number of single men and women. In 1991-92, the number of new marriages fell to 3.22 million, a drop of 7.1 % from the level in 1981-82.

Table 1: Data Summary

A: US Census Data

	1970	1980	1990
Number of Available Males, (mill.)	16.018	23.412	28.417
Percentage change		46.2	21.4
Number of Available Females, (mill.)	19.592	27.225	31.563
Percentage change		39.0	15.9
Average age of Available Males	30.4	29.6	31.7
Average age of Available Females	39.1	37.1	37.9

B: Vital Statistics Data

	1969-71	1979-81	1989-91
Average Number of marriages (mill.)	3.236	3.449	3.220
Percentage change		6.6	-7.11
Average age of Married Males	27.1	29.2	31.2
Average age of Married Females	24.5	26.4	28.9
Average couple age difference	2.6	2.7	2.3

3.2 *Estimating the Gains to Marriage*

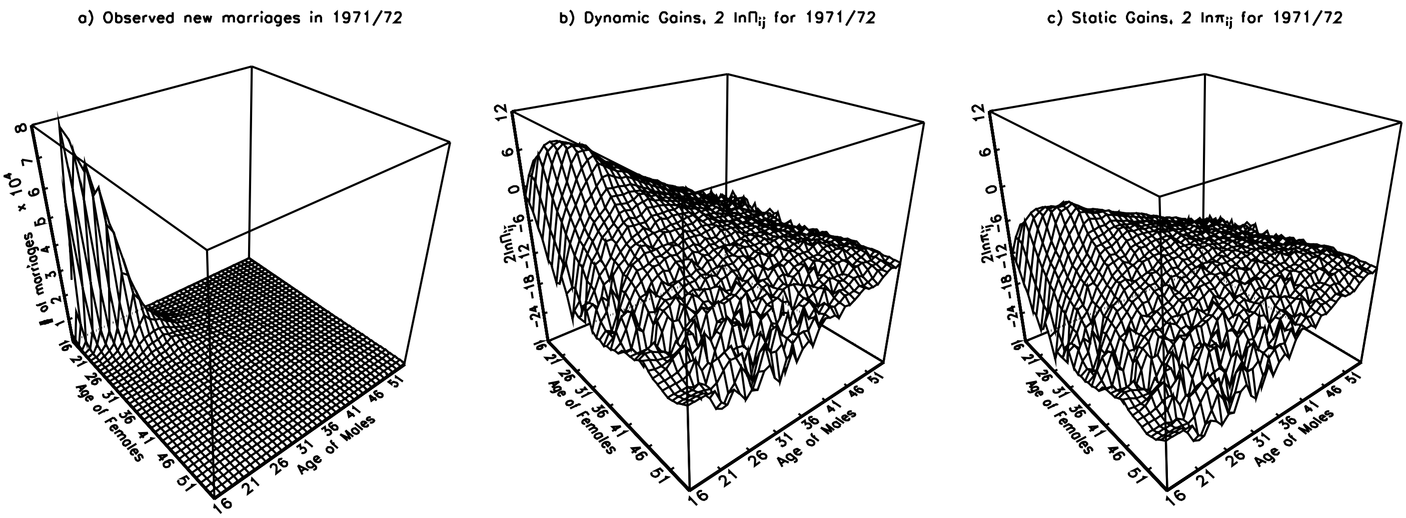

FIGURE 2

Figure 2 a) graphs the distribution of new marriages by age in 1971-72, while Figure 2 b) and 2 c) graphs the estimates of the Static Gains and Dynamic Gains from marriage implied by this distribution of new marriages.[23] The Static Gains plot in Figure 2 b) shows strong assortative matching by age with the gains being highest for couples that are close in age along the diagonal.[24] It shows a peak occurring at an early age when young couples matched with each other. The plot of the Dynamic Gains in Figure 2 b) maintains much of the qualitative features of Figure 2 c). The Dynamic Gains plot strong assortative matching pattern by age with the peak being a lot higher and occurring at an even earlier age. Aside from the difference in the peak, a significant portion of the Dynamic Gains for young age couples are now positive compared to the case where the Static Gains are all negative.

The Static Gains for an (i, j) pair is computed by taking the natural log of the number of current (i, j) matches divided by the geometric averages of those i and j types that

[23] In Figure 2 b) and 2 c), a non-parametric estimate is used to predict the gains for those age pairs where no marriages is observed. This typically happens for matches with large age differential, that is when a young individual is matched with a much older individual.

[24] Choo and Siow (2006) provides a smoothed version of the plot for the Static Gains in Figure 2 c).

chose to remain single. It ignores dynamic considerations in terms of forgone future opportunities of participating in the marriage market if individuals remain single. The Dynamic Gains statistics compensates for this shortcoming by internalizing the future marriage market opportunities forgone in the gains calculation. It approximates the future value of participating in the marriage market using the probabilities of remaining single in the future given by $\prod_{k=0}^{z_{ij}}(p_{i+k,0}q_{0j+k})^{(\beta S)^k}$. Young individuals have the most opportunity to participate in the marriage market, albeit as older individuals as they age. Given that most marriages occur when individuals are young, the implied Dynamic Gains from marriage accounting for the forgone marriage market opportunities in the future is much larger than the Static Gains. The Static Gains statistics in effect assume that there is only one opportunity to match and that in the future, agents would remain single with certainty. In other words future probabilities of remaining single, $p_{i+k,0}$ and $q_{i+k,0}$ equal 1.

Figure 3 graphs various cross-sections of the 1971-72 Static and Dynamic Gains against the age of their spouse on the horizontal axis. Figure 3 a) and 3 b) plots the marriage gains for females aged 18, 25 and 34 years old and Figure 3 c) and 3 d) plots the gains for males for the same ages. The graphs also plot the bootstrap 95% confidence interval computed using the procedure described in Section 2.6.2.[25] These set of graphs provide a more detailed picture of the differences between the Static and Dynamic Gains from marriage. In terms of magnitude, it is clear that the difference between these two statistics is biggest when at least one of the spouses are young. The Dynamic Gains for an 18 and 25 year old far exceed their corresponding Static Gains and is positive when the spouse is young. The tight computed bootstrap confidence interval also suggest that these estimates seem precisely estimated especially for young individuals where most of the data lies.

[25]For age pairs where no matches were observed, a non-parametric conditional mean was used to predict the gains. In those cases, no standard error nor confidence interval is computed accounting for the gaps in the plots of the 95% confidence intervals.

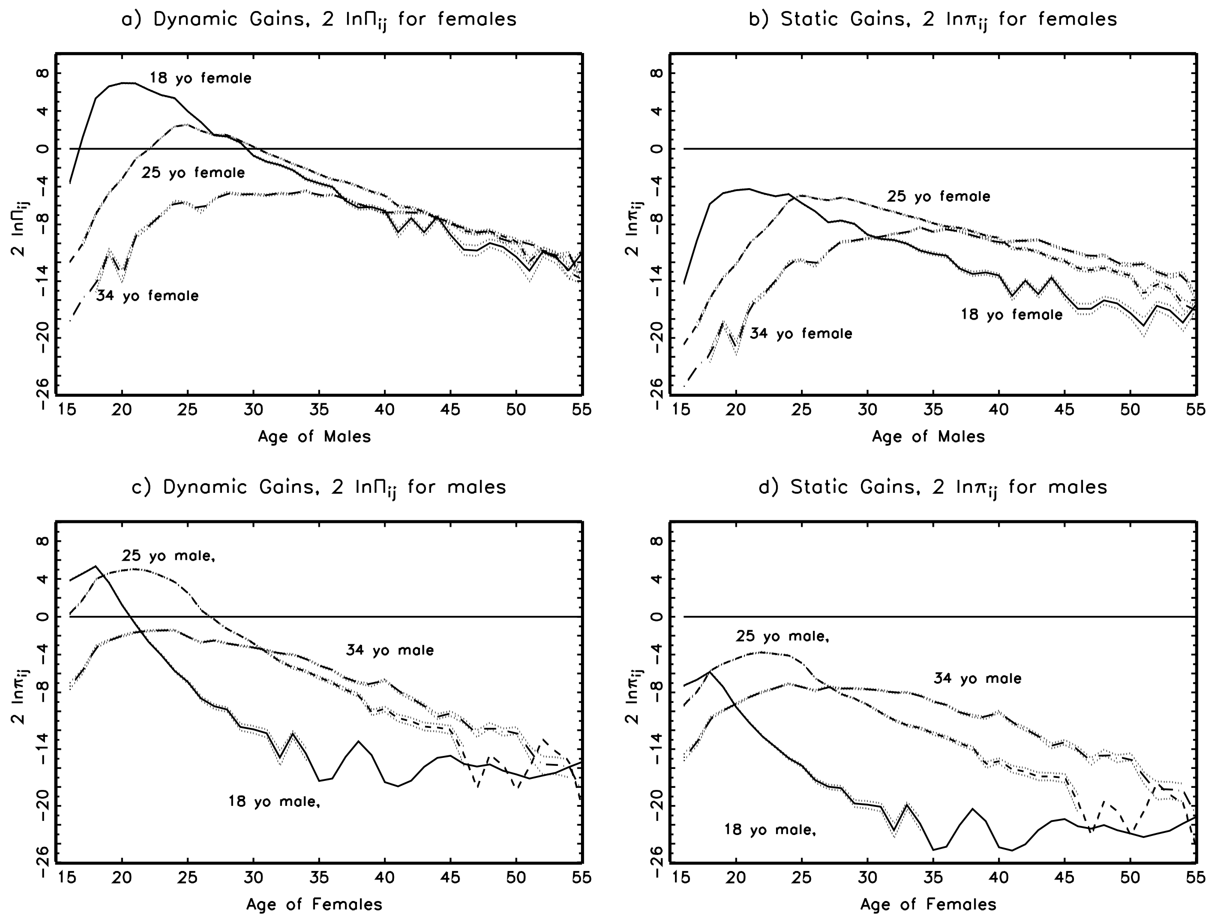

FIGURE 3

To understand how the marriage gains have changed from 1970 to 1980, I construct a simple difference in the Static and Dynamic Gains to marriage. Figure 4 plots various cross sections of this difference against the age of their spouse. Figure 4 a) and 4 b) plots the differences for males and 4 c) and 4 d) plots the differences for females. Generally all these plots suggest that there has been a fall in the gains to marriage over this decade. After accounting for the forgone future marriage opportunities using the Dynamic Gains, Figure 4 a) and 4 c) suggest that the drop in the gains to marriage is even larger than initially suggested by the Static Gains calculation. Comparing Figure 4 a) and b), the drop in marriage gains for 18 years old males are no longer confined to matches with spouses age 18 to 22 but experienced over the entire distribution of spouses' age. While differing in magnitude, the qualitative features of the plots for 25 and 34 year old males are very similar.

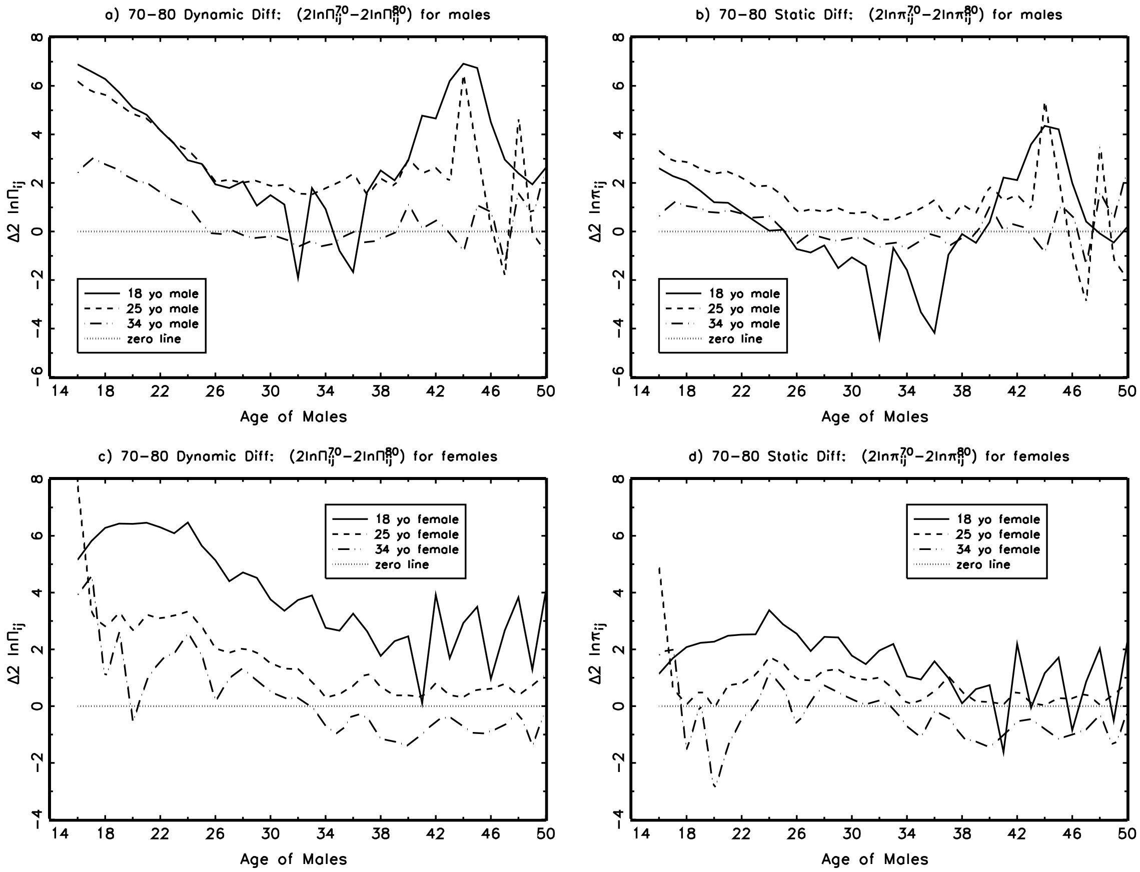

FIGURE 4

4 Conclusion

I propose and estimate a dynamic model of marriage matching. It generalizes the contribution of Choo and Siow (2006) into a dynamic setting while maintaining the empirical tractability and convenience of the static model. Applying the model to US marriage data, I show that ignoring the dynamic returns from marriage severely understates the gains to marriage especially among the young. The proposed framework is sufficiently flexible to allow for matching along other attributes.

References

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

[2] Berry, Steve "Estimating Discrete Choice Models of Product Differentiation," *Rand Journal of Economics*, 25, 242-262.

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

[4] Chiappori, Pierre-Andre, Robert McCann, and Lars Nesheim (2008). "Hedonic Price Equilibria, Stable Matching, and Optimal Transport: Equivalence, Topology, and Uniqueness," Ê̂conomic Theory, 1-49.

[5] Chiappori, Pierre-Andre, Bernard Salaniè, and Yoram Weiss (2011): "Partner Choice and the Marital College Premium," mimeo Columbia University.

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

[7] Choo, Eugene (2012): "Asymptotic Properties of preference estimators in Marriage Matching Models, textquotedblright mimeo University of Calgary.

[8] Debrue, Gerald (1960):"Review of R. D. Luce Individual Choice Behavior" *American Economic Review* 50, 186-188

[9] Decker, Colin, Elliot H. Lieb, Robert McCann, and Benjamin Stephens (2010): "Unique Equilibria and Substitution Effects in a Stochastic Model of the Marriage Market,"mimeo University of Toronto.

[10] Dugar, Subhasish, Haimanti Bhattacharya, and David Reiley (2011) "Can't Buy Me Love? A Field Experiment Exploring the Trade-off Between Income and Caste-Status in an Indian Matrimonial Market", Forthcoming in *Economic Inquiry.*

[11] Echenique, Federico , SangMok Lee, Matt Shum and M. Bumin Yenmez (2012):"The Revealed Preference Theory of Stable and Extremal Stable Matchings" forthcoming *Econometrica.*

[12] Fox, Jeremy T., "Identification in Matching Games," *Quantitative Economics*, 1, 203-254.

[13] Galichon, Alfred, and B. Salaniè (2010): "Matching with Trade-offs: Revealed Preferences over Competing Characteristics," Discussion Paper 7858, CEPR.

[14] Galichon, Alfred, and B. Salaniè (2011): "Cupid's Invisible Hands: Social Surplus and Identification in Matching Models," mimeo.

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

[16] Gretsky, Neil E., Joseph M. Ostroy, and William R. Zame (1999): "Perfect competition in the continuous assignment model" *Journal of Economic Theory*, 88: 60-118.

[17] Hitsch, Günter, Ali Hortaçsu, and Dan Ariely (2010): "Matching and Sorting in Online Dating Markets", 100 (1), 130-163, *American Economic Review*

[18] Lee, Soohyung, and Muriel Niederle, (2011) "Propose with a Rose? Signaling in Internet Dating Markets"mimeo University of Maryland.

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

[20] Rust, John. "Structural Estimation of Markov Decision Processes,"in *Handbook of Econometrics, Volume 4*, ed. by R.E. Engle and D. McFadden. Amsterdam: Elsevier-North Holland, Chapter 14 (1994):

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

[22] Shapley, Lloyd, and Martin Shubik (1972): " The assignment game I: the core," *International Journal of Game Theory*, 1, 111-130.

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

5 Appendix

Lemma 2 *A set of necessary and sufficient conditions for stability is that*

1. *For all matched couples of type (i, j) where male $g \in i$ and female $G \in j$, we require that*

$$\begin{aligned} \epsilon_{ijg} - \epsilon_{ikg} &\geq \tilde{v}_{ik} - \tilde{v}_{ij} \quad \text{for all } k \qquad (32) \\ \epsilon_{ijg} - \epsilon_{i0g} &\geq \tilde{v}_{i0} - \tilde{v}_{ij} \qquad (33) \end{aligned}$$

and

$$\begin{aligned} \epsilon_{ijG} - \epsilon_{kjG} &\geq \tilde{w}_{kj} - \tilde{w}_{ij} \quad \text{for all } k \qquad (34) \\ \epsilon_{ijG} - \epsilon_{0jG} &\geq \tilde{w}_{0j} - \tilde{w}_{ij} \qquad (35) \end{aligned}$$

2. *For every single male g of type i, we require that*

$$\epsilon_{i0g} - \epsilon_{ikg} \geq \tilde{v}_{ij} - \tilde{v}_{i0} \quad \text{for all } k \qquad (36)$$

3. *For every single female G of type j, we require that*

$$\epsilon_{0jG} - \epsilon_{kjG} \geq \tilde{w}_{kj} - \tilde{w}_{0j} \quad \text{for all } k \qquad (37)$$

Proof. See Chiappori, Salaniè and Weiss(2011). ■