Econometrica, Vol. 83, No. 4 (July, 2015), 1373–1423

DYNAMIC MARRIAGE MATCHING: AN EMPIRICAL FRAMEWORK

BY EUGENE CHOO[1]

This paper develops a new model for empirically analyzing dynamic matching in the marriage market and then applies that model to recent changes in the U.S. marriage distribution. Its primary objective is to estimate gains by age from being married today (till death of at least one spouse) relative to remaining single for that same time period. An empirical methodology that relies on the model's equilibrium outcomes identifies the marriage gains using a single cross-section of observed aggregate matches. This behavioral dynamic model rationalizes a new marriage matching function. The model also solves the inverse problem of computing the vector of aggregate marriages, given a new distribution of available single individuals and estimated preferences. Finally, this paper develops a simple test of the model's empirical validity. Using aggregate data of new marriages and available single men and women in the United States over two decades from 1970 to 1990, I investigate the changes in marriage gains over this period.

KEYWORDS: Marriage matching, dynamic marriage, transferable utility.

"Difficult to see. Always in motion is the future."
—Yoda, *Star Wars—The Empire Strikes Back (1980)*

1. INTRODUCTION

THIS PAPER DEVELOPS A MODEL for empirically analyzing dynamic matching in the marriage market and then applies that model to recent changes in the U.S. marriage distribution. Its primary objective is to estimate the dynamic gains by age from being married today (till death of at least one spouse) relative to remaining single (for that same time span). Marriage is inherently fluid with heterogeneous participants, who differ by age, race, income, etc., entering and leaving matches. Unattached individuals can also forgo matching today with the hope of finding a better match in the future, albeit as older, different (e.g., perhaps more educated and wealthier) individuals. The new behavioral overlapping generations model developed in this paper emphasizes these important dynamic features of the marriage market. I develop a new empirical identification strategy that relies on the model's equilibrium outcomes to identify marriage gains using a single cross-section of observed aggregate matches. The proposed methodology is simple to implement in empirical application. The identification strategy of marriage gains is transparent and the identified

[1]I would like to thank a co-editor, three anonymous referees, Liang (Charles) Chen, Kimberly Gilbert, Shashi Khatri, Kevin Laughren, Miguel Olivo-Villabrille, Rob Oxoby, Harry Paarsch, Joanne Roberts, Matt Shum, Scott Taylor, and Arezou Zaresani for their insightful comments. I also thank seminar participants at Singapore Management University. Part of this research was done while on sabbatical at the Economics Department at N.U.S.; 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.

 DOI: 10.3982/ECTA10675

parameters are simple to interpret. These estimates are consistent with a dynamic equilibrium model of frictionless marriage matching with transferable utility. It also rationalizes a new closed-form marriage matching function.[2]

The model extends the first empirical implementation of the static Becker–Shapley–Shubik marriage matching model proposed in Choo and Siow (2006) (CS hereafter) by positing a distribution of male and female preferences over all possible spouse types (which, for expositional convenience, has been restricted to age). In CS, these preferences were explicitly aggregated to derive a quasi-demand and quasi-supply for spouses by age, which in turn are combined with an equilibrium market clearing condition to generate the equilibrium number of marriages. In contrast, the new model in this paper uses the dynamic discrete choice framework of Rust (1987). It rationalizes a new marriage matching function in which, for each age pair, the flow of new marriages is a function of the number of unmatched agents by age (today and in the future) and the intrinsic marriage gains parameters. These parameters are model primitives and invariant to shocks in the number of available single men and women. The marriage gains parameter measures the difference between the discounted within-match utility of being married today (till death of at least one spouse) and that of remaining single (for that same time span). I show that these parameters are identified from the ratio of the product of the equilibrium probabilities of marriage for the couple's age groups, over a weighted average of the equilibrium probabilities that these age groups remain single in the future. The ratio's denominator captures the expected opportunity cost incurred from matching and forgoing future participation in the marriage market. Since marriage might end in divorce, the weights on the probabilities of remaining single reflect the differences in the opportunity costs between the choice to marry and the choice to remain single. The model also solves the inverse problem of computing the vector of aggregate marriage, given a new distribution of available single individuals and estimated preferences. This permits me to compute counterfactual marriage distributions when there are demographic changes, assuming that preferences remain fixed.

Using aggregate data of new marriages and available single men and women in the United States over two decades from 1970 to 1990, I demonstrate the model's application by examining the changes in the marriage gains over these decades and compare the results with those obtained from a static model. During the period from 1970 to 1990, there has been a well-documented fall in the marriage rates in the United States. Part of this decline can be explained by socio-political changes that affected the institution of marriage.[3] I use the

[2]This paper also develops a simple test of the model's empirical validity (in Section 3) and a bootstrap procedure for deriving the standard errors of the estimated preferences (in the Supplemental Material (Choo (2015))).

[3]For instance, some argue that the national legalization of abortion following the U.S. Supreme Court ruling on Roe versus Wade (1973) has lowered the marriage gains.

proposed empirical model to estimate marriage gains over this period and compare them with marriage gains estimates ignoring dynamic considerations. I show that the contribution of the dynamic component of the total marriage gains is large, especially since most marriages occur when individuals are young and face many future opportunities to participate in the marriage market as they age. The decision to marry early suggests that the implied present discounted relative returns from locking into marriage early are high. When analyzing the change in marriage gains over these two decades, I show that ignoring the dynamic component of marriage gains severely understates the decline in marriage gains among the young.

I now briefly describe the main features of the model. The model assumes that the observed aggregate matches by age come from a stationary distribution of available adults with a data generating process that is in steady state. In each period, a constant number of single young men and women enters the system. As in CS, I assume that all individuals of the same age have identical mean preferences and an individual specific additively separable idiosyncratic preference parameter that is independent of the spouse's identity. This ensures enough model heterogeneity in preferences among observationally identical individuals to rationalize observed aggregate matches. In the model, a single unattached individual decides whether to marry or remain single, while taking into account how this decision affects future choices. The choice to marry incurs a one-time transfer of utility at the time of marriage that is dependent only on the ages (i.e., types) of the couple (and not on the identity of the individuals matching). Similarly, the mean joint marriage payoffs depend on the ages of the individuals matching and the expected value of being single in the future should the marriage dissolve. A decision to marry "locks" an individual into a stream of age-dependent, within-marriage utilities, with an exogenous probability of divorce at each age. As an individual ages, the expected value of participating in the marriage market also changes.[4]

For expositional convenience, I first describe the model assuming that (i) divorce occurs at a constant exogenous rate, (ii) the vector of available men and women by age is exogenously given, and (iii) agents live for a fixed time period. In the empirical application, the model is extended to (i) allow for heterogeneity in divorce rates that depends on the couple's age and the tenure of the marriage, (ii) allow the number of single individuals to be endogenously determined through a system of demographic accounting equations, and (iii) allow for differential mortality by gender. Section 4 and the Supplemental Material (Choo (2015)) provide a detailed discussion of these extensions. While the model allows for heterogeneity in the divorce rates, it still assumes that divorce is exogenous. Modeling how divorce occurs more formally is beyond this

[4]The framework can be easily extended to allow for other discrete characteristics, such as race and education.

paper's scope.[5] As in CS, I assume that the additively separable idiosyncratic shock is independent and identically distributed (i.i.d.) with a Type I Extreme Value distribution. The Supplemental Material also considers relaxing this assumption by allowing for correlation across idiosyncratic draws.

The rest of the paper proceeds as follows: Section 2 discusses related papers. I present the model in Section 3. In Section 4, I discuss relaxing some of the model's assumptions and examine the implications. Section 5 describes the empirical application and documents the results. Section 6 concludes.

2. RELATED LITERATURE

The empirical framework presented in this paper owes much of its spirit to the empirical industrial organization literature. Many papers on demand and supply estimation have focused on identifying consumers' preferences and firms' cost parameters from market-level data. Seminal papers such as Bresnahan (1987), Berry (1994), and Berry, Levinsohn, and Pakes (1995), among others, continue the tradition of the discrete choice literature by developing methodologies to estimate individuals' demand preference parameters over attributes. As has become the dominant standard in the demand and supply literature, I assume that agents in my model value the characteristics of the person with whom they are matching.

There is a growing body of structural empirical papers on matching applied to a variety of matching markets. Recent papers include Chen (2013), who looked at matching of CEOs and firms; Agrawal (2013), who looked at placement of medical residents; and Fox (2010), who proposed a maximum-score estimator to analyze firm-level car parts matching data. A number of papers have proposed generalizations to the CS empirical framework. These include Galichon and Salanié (2012), Chiappori, Salanié, and Weiss (2011), Graham (2011, 2013). Decker, Lieb, McCann, and Stephens (2013) provided a test for the CS model that exploits the symmetry restrictions on the cross-type marriage elasticity matrix.[6]

Few papers apply non-transferable utility models to individual-level matchings. Echenique, Lee, Shum, and Yenmez (2013) is the first paper to derive

[5]This is an important area of research that has long interested social scientists. Modeling how match values evolve would, however, require more detailed individual-level within-marriage data that are not used in this paper. See Brien, Lillard, and Stern (2006).

[6]The symmetry restriction requires that the elasticity of type i single men with respect to the supply of type j women be equal to the elasticity of type j single women with respect to the supply of type i men. This restriction is reminiscent of the Independence of Irrelevant Alternatives (I.I.A.) property brought about by the i.i.d. additive utility error imposed by the discrete choice structure (see McFadden (1974) and Debreu (1960)). The exact cost of this restriction in the context of marriage matching models that maintain the CS structure remains to be seen. A large body of papers in Empirical Industrial Organization and the discrete choice literature is devoted to overcoming the I.I.A. properties of the discrete choice models. Some papers include McFadden (1978), Berry, Levinsohn, and Pakes (1995), and Petrin (2002).

the empirical implications of stable two-sided matching. Menzel (2015) studied two-sided matching markets with non-transferable utility when the number of market participants grows large. The author derived tractable asymptotic approximation to the distribution of matched observable characteristics resulting from pairwise stable matchings. Hsieh (2013) separately identified male and female marital preferences using a non-transferable utility model applied to aggregate marriage data. A related paper using the non-transferable utility specification is Hitsch, Hortaçsu, and Ariely (2010), which focused on identifying preferences separately from the matching process. Employing a data set from an online dating service, the authors estimated a rich specification of preference over spousal physical and socio-economic characteristics.[7]

Many researchers have recognized that the marriage market is an important determinant of redistribution between men and women. Many empirical studies also suggest there may be strong connections between features of the marriage market, such as sex ratios and divorce laws, and redistribution within marriage and household labor supply behavior. (See, e.g., Chiappori, Fortin, and Lacroix (2002), Amuedo-Dorantes and Grossbard-Schectman (2007), and Lundberg and Pollak (1996), among others.) These studies motivate empirical investigation of the joint determination of marriage matching and intra-household allocations. Choo and Seitz (2013) developed the collective marriage matching model that integrates the collective framework of intra-household decision making of Chiappori (1988, 1992) and the marriage matching model of CS. The authors' framework of endogenizing marriage decisions in the collective model generates two independent sets of sharing rule estimates: one from labor supplies and one from marriage decisions.[8] The papers discussed so far all focused on static one-shot matching. The idea of modeling marriage decisions using a dynamic structural model is not new and several papers estimate the model primitives using longitudinal data. These papers include van der Klaauw (1996), Seitz (2009), Brien, Lillard, and Stern (2006), and Bruze, Svarer, and Weiss (2015) among others.[9]

[7]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 increase chances in online dating.

[8]The collective marriage matching model in Choo and Seitz (2013) is based on the collective model of Chiappori, Fortin, and Lacroix (2002). As in Chiappori, Fortin, and Lacroix (2002), Choo and Seitz (2013) focused on households in which all consumption and leisure are private. In such a setting, the household's Pareto problem can be decentralized and the resource allocation in the household can be summarized by a lump sum transfer of income, the sharing rule. Choo and Seitz (2013) produced a sharing rule that is the equilibrium outcome of marriage matching. The authors showed that partial derivatives of the equilibrium sharing rules are identified from marriage data.

[9]An earlier paper, Choo and Siow (2007), shares a similar objective to the current paper in its attempt to model the bivariate marriage distribution by age using the dynamic discrete choice framework of Rust (1987). It focused on developing a general equilibrium framework to test different theories of marital and home production. The framework puts restrictions on linearized

3. A DYNAMIC MATCHING MODEL

3.1. *Preview of Results*

Consider a stationary environment in which men and women live for Z periods. The equilibrium matching model delivers a new marriage matching function:[10]

$$\mu_{i,j} = \widehat{\Pi}_{i,j}\sqrt{m_i f_j}\prod_{k=0}^{T_{i,j}-1}\left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k}f_{j+k}}\right)^{(1/2)(\beta(1-\delta))^k}. \tag{3.1}$$

Equation (3.1) expresses the number of new marriages between age i men and age j women, $\mu_{i,j}$, as a function of the model primitives estimate, $\widehat{\Pi}_{i,j}$; the stationary number of available age i men and available age j women, m_i and f_j, respectively; and a "scaled average" of the proportion of single age i men, $(\mu_{i,0}/m_i)$, and single age j women, $(\mu_{0,j}/f_j)$. The term $T_{i,j} = Z - \max(i, j) + 1 \geq 1$ represents the maximum length of a match before one of the spouses in the match passes away at the terminal age Z. For convenience, I refer to $T_{i,j}$ as *the duration of the match*. $(1 - \delta)$ denotes the probability a marriage survives in any period and β is the per-period discount factor.[11] $\mu_{i,0}$ and $\mu_{0,j}$ denote the number of type i men and type j women who chose to remain single at age i and j, respectively. The parameter $\widehat{\Pi}_{i,j}$ represents an estimate of marriage gains from an (i, j) marriage relative to remaining single for the duration of the match.

As in CS, the matching function (3.1) is homogeneous of degree zero in the vector of single individuals and marriages and allows for substitution effects across all spouse ages.[12] That is, holding all else constant, a change in $m_{i'}$ or

per-period marriage gains. The linearization in Choo and Siow (2007) allows the authors to empirically approximate the benefit of delaying marriage for one period versus marrying today using the marriage growth rate. This approximation then becomes the basis on which the authors test different theories of marital and home production. While the present paper shares many similar objectives with Choo and Siow (2007), the representation of the dynamic problem proposed in the present paper is new. This new representation permits me to derive a closed-form marriage matching function that is the dynamic analogue of the one proposed in CS. The estimation of marital surplus developed in this paper using a single cross-section of aggregate matches is simple and transparent.

[10]A *marriage matching function*, denoted by $\boldsymbol{\mu} = \mathcal{G}(\mathbf{m}, \mathbf{f}; \boldsymbol{\Pi})$, is a simple reduced-form way of characterizing the entire distribution of aggregate marriages by age, denoted by a $(Z \times Z)$ matrix of marriages $\boldsymbol{\mu}$ as a function of exogenous factors that include a $(Z \times 1)$ vector of available single men, $\mathbf{m}$, a $(Z \times 1)$ vector of available single women, $\mathbf{f}$, and a $(Z \times Z)$ matrix of parameters, $\boldsymbol{\Pi}$. $\mu_{i,j}$ is the (i, j)th element of $\boldsymbol{\mu}$, m_i the ith element of $\mathbf{m}$, and f_j the jth element of $\mathbf{f}$, and $\Pi_{i,j}$ denotes the (i, j)th element of the matrix $\boldsymbol{\Pi}$.

[11]For ease of exposition, the survival rate (which is 1 less the divorce rate) of a marriage is assumed to be constant across ages of matches and marriage tenure. Section 4 relaxes this assumption by allowing heterogeneity in the divorce rate.

[12]The CS marriage matching function is $\mu_{i,j} = \pi_{i,j}\sqrt{\mu_{i,0}\mu_{0,j}}$.

$f_{j'}$ will lead to a change in the number of equilibrium (i, j) matches, $\mu_{i,j}$, even if $i' \neq i$ and $j' \neq j$. Another unique feature of this function is the interdependence between the number of equilibrium (i, j) matches and their respective future marriage prospects as captured by the equilibrium probabilities of remaining single in the future, $(\mu_{i+k,0}/m_{i+k})$ and $(\mu_{0,j+k}/f_{j+k})$. This interdependence arises because of the dynamic feature in the matching model.[13]

Moving observed quantities to one side and model parameters to the other and taking logs, I obtain

$$2\ln\widehat{\Pi}_{i,j} = \underbrace{\left[\ln\frac{\mu_{i,j}}{m_i} - \sum_{k=0}^{T_{i,j}-1}\ln\left(\frac{\mu_{i+k,0}}{m_{i+k}}\right)^{(\beta(1-\delta))^k}\right]}_{n_{i,j}(\cdot)} + \underbrace{\left[\ln\frac{\mu_{i,j}}{f_j} - \sum_{k=0}^{T_{i,j}-1}\ln\left(\frac{\mu_{0,j+k}}{f_{j+k}}\right)^{(\beta(1-\delta))^k}\right]}_{\mathcal{N}_{i,j}(\cdot)}. \tag{3.2}$$

This marital surplus captured by $2\ln\widehat{\Pi}_{i,j}$ represents the couple's discounted within-marriage utilities from being locked in an (i, j) match today for the duration of the match relative to the present discounted sum of per-period payoffs for both individuals remaining single for the entire period. It comprises the male and female net gains from marriage relative to being single for the duration of the match as represented by $n_{i,j}(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f})$ and $\mathcal{N}_{i,j}(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f})$, respectively. The tuple $(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f})$ represents the matrix of observed marriages and the vector of available men and women, respectively.[14]

The model assumptions (in particular, the Type 1 Extreme Value distributional assumption together with the transferable utility setup) allow me to express the net marital surplus as a *dynamic log-odds ratio*. This dynamic analogue to a "static" log-odds ratio is composed of an age i male's equilibrium

[13]This marriage matching function also needs to satisfy a set of accounting constraints. The accounting constraints are

$$\mu_{0,j} + \sum_{i=1}^{Z}\mu_{i,j} = f_j \quad \forall j,$$

$$\mu_{i,0} + \sum_{j=1}^{Z}\mu_{i,j} = m_i \quad \forall i,$$

$$\mu_{0,j}, \mu_{i,0}, \mu_{i,j} \geq 0 \quad \text{for all } i \text{ and } j.$$

I will describe these constraints in more detail in Section 3.3.

[14]Both $n_{i,j}(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f})$ and $\mathcal{N}_{i,j}(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f})$ are endogenous quantities which depend on the relative scarcity of available men and women.

probability of an (i, j) marriage as estimated by $(\mu_{i,j}/m_i)$ over a weighted average of the equilibrium probabilities that an age i male would remain single in the future as estimated by $\prod_{k=0}^{T_{i,j}-1}(\mu_{i+k,0}/m_{i+k})^{(\beta(1-\delta))^k}$. The equilibrium probability of remaining single at a future age captures the expected value of being single and participating in the marriage market at that age. In the event that an (i, j) marriage survives for k periods, the age i male's expected cost of being unable to participate in the marriage market when he is age $i+k$ is captured by $(\mu_{i+k,0}/m_{i+k})$. If the expected marriage rates at $i+k$ are high, then the opportunity cost for men of being locked in a marriage at age $i+k$, rather than being able to participate in the marriage market, is also high. This result implies that male net gains from the choice of an (i, j) marriage relative to staying single are large.[15] The interpretation of $\mathcal{N}_{i,j}(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f})$ is similar.

Equation (3.2) is also the estimating equation for the parameters $2\ln\widehat{\Pi}_{ij}$. 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. In empirical application, practitioners are often interested in the *inverse* problem. That is, given an estimated vector $\widehat{\boldsymbol{\Pi}}$ consistent with a vector of aggregate matches $(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f})$, satisfying Equation (3.1) and the accounting constraints, how do changes in the vector of available men and women affect the distribution of matches? I use this model to analyze the change in marriage gains in the United States between 1970 and 1990 and compare these results with the static matching model of CS.

3.2. *The Model Environment*

Consider a stationary society in discrete time populated by overlapping generations of adults. Each individual lives for Z periods irrespective of gender.[16] 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. Agents are horizontally differentiated by age only.[17] m_i and f_j denote the numbers of single males of age i and females of

[15]In other words, when comparing male net marriage gains for pairs (i, j) and (i', j') that have the same duration of marriage and the same male probabilities of marriage (i.e., $(\mu_{i,j}/m_i) = (\mu_{i',j'}/m_{i'})$), if the future marriage rates for i are greater than i', I expect the male net gains from an (i, j) match to be larger than those of (i', j'). The reason is that in choosing to marry age j women, age i men, on average, are incurring a greater opportunity cost than age i' men who marry age j' women (i.e., $n_{i,j}(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f}) > n_{i',j'}(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f})$). Since a decision by the couple to marry at ages i and j does not completely preclude the couple from participating again in the marriage market in the future, the weights in the ratio reflect the difference in the forgone future opportunity from marriage relative to being single.

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

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

age j at the beginning of each period and I assume the vectors of single men and women $\mathbf{m} = (m_1, \dots, m_Z)'$ and $\mathbf{f} = (f_1, \dots, f_Z)'$ as exogenous and given. I relax this assumption in Section 4.

Any *single* age i male g (or age j female h) in each period is characterized by two *state variables*:

- i (or j) $\in \{1, \dots, Z\}$ is his (or her) age when single, and
- $\boldsymbol{\varepsilon}_{i,g}$ (or $\boldsymbol{\varepsilon}_{j,h}$) is a $(Z+1)$ vector of i.i.d. idiosyncratic payoffs or match-specific errors specific to age i male, g (or age j female, h) that are unobserved to the econometrician.

At each period, a single age i male (or single age j female) faces a random utility draw from each type of spouse available and from remaining single. He or she chooses the option maximizing his or her discounted expected utility. Male g observes the $(Z+1) \times 1$ vector of idiosyncratic payoffs $\boldsymbol{\varepsilon}_{i,g} = (\varepsilon_{i,0,g}, \varepsilon_{i,1,g}, \dots, \varepsilon_{i,Z,g})'$ at the beginning of each period before deciding on a utility maximizing decision. $\varepsilon_{i,0,g}$ is the idiosyncratic payoff that g receives from remaining single, and $\varepsilon_{i,j,g}$ is the idiosyncratic payoff that g receives from matching with an age j female. Similarly, female h observes a $(Z+1) \times 1$ vector of idiosyncratic payoffs $\boldsymbol{\varepsilon}_{j,h} = (\varepsilon_{0,j,h}, \varepsilon_{1,j,h}, \dots, \varepsilon_{Z,j,h})'$.

These type-specific idiosyncratic draws do not depend on the identity of the spouse the single decision maker meets or matches with; they depend only on the identity of the decision maker male g (or female h). The crucial assumption first introduced in CS provides empirical tractability in explaining observed aggregate matches. In other words, the idiosyncratic contributions of two observationally equivalent age j female partners to male g's utility are the same. I assume that the i.i.d. $\boldsymbol{\varepsilon}$ are drawn from McFadden's Type 1 Extreme Value distribution.[18] In the Supplemental Material, I consider $\boldsymbol{\varepsilon}$ drawn from the family of Generalized Extreme Value distributions to allow for persistence in the idiosyncratic preferences.

Only single adults can make utility maximizing decisions whether or not to marry. A decision to marry locks an individual into a stream of payoffs in the event the marriage does not dissolve due to divorce or death. Marriages dissolve at a constant exogenous per-period rate, δ, for all (i, j) pairs.[19] Divorce shocks are realized just before the end of a period after utility maximizing decisions are made. A single individual who chooses to marry during a period could be divorced before that period ends, in which case he (or she) re-enters the marriage market in the next period as a one-period-older individual. If divorce occurs in period k of a marriage, where $1 \le k \le (Z - \max(i, j) + 1)$, the individuals g and h re-enter the marriage market as single individuals of age $i + k$ and

[18] The marginal density is given by $\mathfrak{f}(\boldsymbol{\varepsilon}_{i,g} \mid i) = \prod_{a=0}^{Z} \exp[-\varepsilon_{i,a,g}] \exp[-\exp(-\varepsilon_{i,a,g})]$.

[19] This approximation, while clearly restrictive, simplifies the equilibrium condition which I take to data and the estimation strategy used to estimate the model primitives. Understanding how within-match value evolves and endogenizing divorce are important topics of research that are beyond the scope of the present paper. In Section 4.2, I formulate a model that allows for heterogeneity in the divorce rate that depends on the couple's ages and the marriage tenure.

$j+k$, respectively. The duration of an (i,j) match, $T_{i,j} = (Z - \max(i,j) + 1)$, is the maximum length of a marriage that ends with the death of the oldest spouse.

Let $\alpha_{i,j,k}$ (or $\gamma_{i,j,k}$) be the kth period within-marriage surplus accrued to an age i male (or j female) when married to an age j female (or age i male) where $k \geq 1$. $a_{i,g}$ (or $a_{j,h}$) denotes the action of a single age i male g (or j female h) where $a_{i,g}$ (or $a_{j,h}$) $\in \mathcal{D} = \{0, 1, \dots, Z\}$. If he (or she) chooses to remain single, $a_{i,g} = 0$ (or $a_{j,h} = 0$). If he (or she) chooses to match with an aged k spouse, $a_{i,g} = k$ (or $a_{j,h} = k$). The one-period utility of male g with state vector $(i, \boldsymbol{\varepsilon}_{i,g})$ and action $a_{i,g}$ is denoted by $v(a_{i,g}, i, \boldsymbol{\varepsilon}_{i,g})$. The utility of female h with state vector $(j, \boldsymbol{\varepsilon}_{j,h})$ and action $a_{j,h}$ is denoted by $w(a_{j,h}, j, \boldsymbol{\varepsilon}_{j,h})$:

$$v(a_{i,g} = j, i, \boldsymbol{\varepsilon}_{i,g}) = \begin{cases} \boldsymbol{\alpha}_i(j) - \tau_{i,j} + \varepsilon_{i,j,g}, & \text{if } a_{i,g} \in \{1, \dots, Z\}, \\ \alpha_{i,0} + \varepsilon_{i,0,g}, & \text{if } a_{i,g} = 0, \end{cases} \tag{3.3}$$

and

$$w(a_{j,h} = i, j, \boldsymbol{\varepsilon}_{j,h}) = \begin{cases} \boldsymbol{\gamma}_j(i) + \tau_{i,j} + \varepsilon_{i,j,h}, & \text{if } a_{j,h} \in \{1, \dots, Z\}, \\ \gamma_{0,j} + \varepsilon_{0,j,h}, & \text{if } a_{j,h} = 0, \end{cases} \tag{3.4}$$

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

$$\boldsymbol{\alpha}_i(j) = \sum_{k=1}^{T_{i,j}} \big(\beta(1-\delta)\big)^{k-1} \alpha_{i,j,k}, \quad \text{and} \quad \boldsymbol{\gamma}_j(i) = \sum_{k=1}^{T_{i,j}} \big(\beta(1-\delta)\big)^{k-1} \gamma_{i,j,k}. \tag{3.5}$$

$\alpha_{i,0}$ and $\gamma_{0,j}$ are the per-period utilities from remaining single for i age males and j age females, respectively. The time discount factor is denoted by $\beta \in (0,1)$.

Equation (3.3) states that if age i male g marries an age j woman, he receives the mean utility from the match equal to $\boldsymbol{\alpha}_i(j) - \tau_{i,j}$, plus an idiosyncratic shock $\varepsilon_{i,j,g}$, with the mean utility from marriage depending only on the age of the men and women in the match.[20] He also receives an idiosyncratic return specific to him, $\varepsilon_{i,j,g}$, that does not depend on the precise identity of g's spouse. The mean utility from marriage comprises two terms. $\boldsymbol{\alpha}_i(j) = \sum_{k=1}^{T_{i,j}} (\beta(1-\delta))^{k-1} \alpha_{i,j,k}$ captures the discounted stream of male within-marriage payoffs in the event that the marriage does not dissolve. If divorce happens in the kth period of marriage, the couple will re-enter the marriage market as an $i+k$ and a $j+k$ age single male and female (assuming that $T_{i,j} \geq k$). If g chooses to remain single, his current-period mean utility (common to all men of his age) is $\alpha_{i,0}$. In choosing this match, g commits to pay a one-off transfer, $\tau_{i,j}$, specific to the age pair (i,j) of individuals matching.

[20]This is, in part, due to the AS assumption, but transferable utility and frictionless matching are also necessary for this.

Similarly, in Equation (3.4), an age j female h who decides to marry an age i man agrees to receive this equilibrium transfer. In accepting the match, she locks herself into a stream of marital payoffs, of which the present discounted value equals $\sum_{k=1}^{T_{i,j}}(\beta(1-\delta))^{k-1}\gamma_{i,j,k}$. Thus, if male g wants to marry female h, he has to transfer $\tau_{i,j}$ of marital output to her. Similarly, if an age j woman h wants to marry an age i man g, she has to be willing to accept $\tau_{i,j}$ of marital output from him. Each individual takes $\tau_{i,j}$ as exogenous. The marriage market clears when, given $\tau_{i,j}$, for every i, j, the number of age i men who want to marry age j women is equal to the number of age j women who want to marry age i men. This transfer can be positive or negative. In this full-commitment model, the one-time payment of $\tau_{i,j}$ fully internalizes the discounted stream of within-marriage utilities for this (i, j) couple, the exogenous divorce probabilities, and the relative scarcity of males and females in the system.

The full-commitment model adopts the standard Beckerian assumption that prospective spouses who marry make binding agreement on the stream of allocations within marriage and the one-off transfer that is specific to the age (or types) of the couples. This transferable utility assumption provides a simple way for each partner to be compensated for the desirability of his or her age (or type) using match-specific marriage gains without affecting aggregate marriage surplus (see Weiss (1997)).[21]

The specification of preferences over partners and the evolution of the state variables satisfy two assumptions: the *Additive Separability* (AS) and *Conditional Independence* (CI) assumptions. Both these assumptions were introduced by Rust (1987) in the context of a single agent dynamic discrete choice model. I assume that the following assumptions hold:

ASSUMPTION AS—Additive Separability: *The utility functions* $v(a_{i,g}, i, \boldsymbol{\varepsilon}_{i,g})$ *and* $w(a_{j,h}, j, \boldsymbol{\varepsilon}_{j,h})$ *have additively separable decompositions of the form*

$$v(a_{i,g}, i, \boldsymbol{\varepsilon}_{i,g}) = v_a(i) + \varepsilon_{i,a,g}, \tag{3.6}$$

$$w(a_{j,h}, j, \boldsymbol{\varepsilon}_{j,h}) = w_a(j) + \varepsilon_{j,a,h}, \tag{3.7}$$

where $\varepsilon_{i,a,g}$ *and* $\varepsilon_{j,a,h}$ *are the ath component of the vector* $\boldsymbol{\varepsilon}_{i,g}$ *and* $\boldsymbol{\varepsilon}_{j,h}$, *respectively.* $v_a(i)$ (*or* $w_a(j)$) *is the mean utility accrued for all age i men* (*or j women*) *who chose action* a. *From Equations* (3.3) *and* (3.4) *above*:

$$v_a(i) = \begin{cases} \boldsymbol{\alpha}_i(j) - \tau_{i,j}, & \text{if } a = j \in \{1, \dots, Z\}, \\ \alpha_{i,0}, & \text{if } a = 0, \end{cases}$$

[21]Given this is a full-commitment model of marriage with exogenous divorce, transfers could equivalently be made as a stream over the course of the marriage. I use a one-time payment for analytical convenience and assume there is no resource or income constraint at the time of marriage.

and

$$w_a(i) = \begin{cases} \boldsymbol{\gamma}_j(i) + \tau_{i,j}, & \text{if } a = j \in \{1, \dots, Z\}, \\ \gamma_{0,j}, & \text{if } a = 0. \end{cases}$$

ASSUMPTION CI—Conditional Independence: *The transition probabilities of the state variables for males and females respectively factorize as*

$$\mathbb{P}\{i', \boldsymbol{\varepsilon}'_{i,g} \mid i, \boldsymbol{\varepsilon}, a\} = \mathfrak{f}(\boldsymbol{\varepsilon}' \mid i) \cdot \mathcal{F}_a(i' \mid i), \tag{3.8}$$

$$\mathbb{P}\{j', \boldsymbol{\varepsilon}'_{j,h} \mid j, \boldsymbol{\varepsilon}, a\} = \mathfrak{f}(\boldsymbol{\varepsilon}' \mid j) \cdot \mathcal{R}_a(j' \mid j), \tag{3.9}$$

where $\mathfrak{f}(\boldsymbol{\varepsilon})$ *is the multivariate p.d.f. of the i.i.d.* $\boldsymbol{\varepsilon}$, *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).

The AS assumption is standard in the static and dynamic discrete choice literature. This structure ensures that the utility function depends separately on a function of observed state and action, $v_a(i)$ and $w_a(j)$, and the individual-specific idiosyncratic taste parameters, $\varepsilon_{i,a,g}$ and $\varepsilon_{a,j,h}$. This restriction on preferences precludes any interaction between the observed state and choice decision, and the idiosyncratic taste term. The CI assumption limits the dependence structure on the state variables. Rust (1994) argued that the observed states i' (and j') are sufficient statistics for the unobserved states $\boldsymbol{\varepsilon}_{i',g}$ (and $\boldsymbol{\varepsilon}_{j',h}$). So any dependence between, say, $\boldsymbol{\varepsilon}_{i,g}$ and $\boldsymbol{\varepsilon}_{i+1,g}$ is only transmitted through the observed age, $i+1$, and not through the unobserved $\boldsymbol{\varepsilon}_{i,g}$. More importantly, the probability that any individual g is single again at age i' depends only on his action a and age i and not on the unobserved state $\boldsymbol{\varepsilon}_{i,g}$. I use $\mathcal{F}_a(i' \mid i)$ (or $\mathcal{R}_a(j' \mid j)$) to denote the transition probability that an age i male g (or j female h) will next find himself (or herself) single at age i' (or j') given his (or her) action a at age i (or j). $\boldsymbol{\varepsilon}$ are i.i.d. noises that are superimposed on this process. Appendix A.1 provides details about the state transition probabilities, $\mathcal{F}_a(i' \mid i)$ and $\mathcal{R}_a(j' \mid j)$.

Figure 1 provides an overview of the sequence of events and decisions for an unattached individual in a given period. A single agent observes the choice-

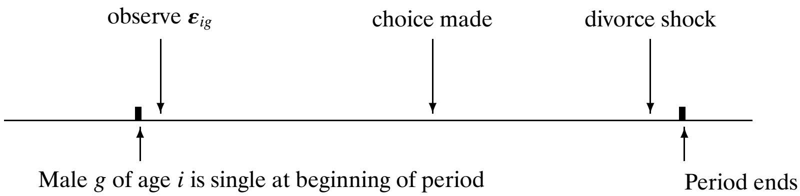

FIGURE 1.—Timing sequence of shocks and decisions.

specific idiosyncratic shocks $\boldsymbol{\varepsilon}$ at the beginning of the period. This is followed by expected utility maximizing decisions. If the agent chooses to marry, a divorce shock is revealed before the period ends.

Since the decision problem for a single individual is a finite-horizon problem, I solve it by backward induction. The model permits agents to make choices only when they are single. Consider a single male g in his terminal age Z; his value function is given by

$$V_{\boldsymbol{\alpha}}(Z, \boldsymbol{\varepsilon}_{i,g}) = \max\{v(0, Z, \boldsymbol{\varepsilon}_{Z,g}), v(1, Z, \boldsymbol{\varepsilon}_{Z,g}), \ldots, v(Z, Z, \boldsymbol{\varepsilon}_{Z,g})\},$$

where the functional form of the utility from choosing action j with state vector $(Z, \boldsymbol{\varepsilon}_{Z,g})$, $v(j, Z, \boldsymbol{\varepsilon}_{Z,g})$, is given by Equation (3.3). Working backward, for $i < Z$, we get the following Bellman equation for a single age i individual g:

$$\begin{aligned}
(3.10)\quad & V_{\boldsymbol{\alpha}}(i, \boldsymbol{\varepsilon}_{i,g}) \\
&= \max\Bigg\{\alpha_{i,0} + \beta\mathbb{E}\big[V_{\boldsymbol{\alpha}}(i+1, \boldsymbol{\varepsilon}_{i+1,g}) \mid i, \boldsymbol{\varepsilon}_{i,g}, a_{i,g} = 0\big] + \varepsilon_{i,0,g}, \\
&\quad \max_{a\in\{1,\ldots,Z\}}\Bigg\{\boldsymbol{\alpha}_i(a) - \tau_{i,a} \\
&\quad + \sum_{k=i+1}^{\min(Z,i+T_{i,a})} \beta^{k-i}\mathbb{E}\big[V_{\boldsymbol{\alpha}}(k, \boldsymbol{\varepsilon}_{k,g}) \mid i, \boldsymbol{\varepsilon}_{i,g}, a\big] + \varepsilon_{i,a,g}\Bigg\}\Bigg\}.
\end{aligned}$$

The CI assumption allows me to factorize the future expectation of the value function from being single at age k (conditional on being single at age i, observing the vector of idiosyncratic payoffs $\boldsymbol{\varepsilon}_{i,g}$ and choosing decision $a_{i,g}$) as

$$\mathbb{E}\big[V_{\boldsymbol{\alpha}}(k, \boldsymbol{\varepsilon}_{k,g}) \mid i, \boldsymbol{\varepsilon}_{i,g}, a_{i,g}\big] = \mathcal{F}_a(k \mid i) \int V_{\boldsymbol{\alpha}}(k, \boldsymbol{\varepsilon}_{k,g})\mathfrak{f}(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon}_g.$$

The latter term, $\int V_{\boldsymbol{\alpha}}(k, \boldsymbol{\varepsilon}_{k,g})\mathfrak{f}(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon}_g$, is referred to as the *integrated value function*. The Type 1 Extreme Value distributional assumption on $\boldsymbol{\varepsilon}_g$ allows it to have a closed-form representation. I will discuss this further below. The Bellman equation for a single age j female (where $j < Z$) takes a similar form,

$$\begin{aligned}
(3.11)\quad & W_{\boldsymbol{\gamma}}(j, \boldsymbol{\varepsilon}_{j,h}) \\
&= \max\Bigg\{\gamma_{0,j} + \beta\mathbb{E}\big[W_{\boldsymbol{\gamma}}(j+1, \boldsymbol{\varepsilon}_{j+1,h}) \mid j, \boldsymbol{\varepsilon}_{j,h}, a_{j,h} = 0\big] + \varepsilon_{0,j,h},
\end{aligned}$$

$$\max_{a\in\{1,\ldots,Z\}}\left\{\boldsymbol{\gamma}_j(a)+\tau_{a,j}\right.$$
$$\left.+\sum_{k=j+1}^{\min(Z,j+T_{a,j})}\beta^{k-j}\mathbb{E}\big[W_{\boldsymbol{\gamma}}(k,\boldsymbol{\varepsilon}_{k,h})\mid j,\boldsymbol{\varepsilon}_{j,h},a\big]+\varepsilon_{a,j,h}\right\}\Bigg\}.$$

Consider decomposing the value functions $V_{\boldsymbol{\alpha}}(i,\boldsymbol{\varepsilon}_{i,g})$ and $W_{\boldsymbol{\gamma}}(j,\boldsymbol{\varepsilon}_{j,h})$ in Equations (3.10) and (3.11) into two parts: (1) a mean component that is dependent on the utility maximizing choice and the current age, and (2) an idiosyncratic component. The mean component is common to all individuals of the same age choosing the same action. $\mathbb{1}(a\neq 0)$ and $\mathbb{1}(a=0)$ are indicator variables for the decisions to marry and to remain single, respectively. I denote the mean component for age i males and j females by $\tilde{v}_{i,j}$ and $\tilde{w}_{i,j}$, respectively:

$$\tag{3.12}\tilde{w}_{i,j}=\begin{cases}\boldsymbol{\gamma}_j(i)+\tau_{i,j}+\displaystyle\sum_{k=j+1}^{\min(Z,j+T_{i,j})}\beta^{k-j}\mathbb{E}\big[W_{\boldsymbol{\gamma}}(k,\boldsymbol{\varepsilon}_{k,h})\mid j,\boldsymbol{\varepsilon}_{j,h},i\big], & \text{if } j<Z,\ i\neq 0,\\ \gamma_{0,j}+\beta\mathbb{E}\big[W_{\boldsymbol{\gamma}}(j+1,\boldsymbol{\varepsilon}_{j+1,h})\mid j,\boldsymbol{\varepsilon}_{j,h},0\big], & \text{if } j<Z,\ i=0,\\ \big[\boldsymbol{\gamma}_Z(i)+\tau_{i,Z}\big]\cdot\mathbb{1}(i\neq 0)+\gamma_{0,Z}\cdot\mathbb{1}(i=0), & \text{if } j=Z,\end{cases}$$

$$\tag{3.13}\tilde{v}_{i,j}=\begin{cases}\boldsymbol{\alpha}_i(j)-\tau_{i,j}+\displaystyle\sum_{k=i+1}^{\min(Z,i+T_{i,j})}\beta^{k-i}\mathbb{E}\big[V_{\boldsymbol{\alpha}}(k,\boldsymbol{\varepsilon}_{k,g})\mid i,\boldsymbol{\varepsilon}_{i,g},j\big], & \text{if } i<Z,\ j\neq 0,\\ \alpha_{i,0}+\beta\mathbb{E}\big[V_{\boldsymbol{\alpha}}(i+1,\boldsymbol{\varepsilon}_{i+1,g})\mid i,\boldsymbol{\varepsilon}_{i,g},0\big], & \text{if } i<Z,\ j=0,\\ \big[\boldsymbol{\alpha}_Z(j)-\tau_{Z,j}\big]\cdot\mathbb{1}(j\neq 0)+\alpha_{0,Z}\cdot\mathbb{1}(j=0), & \text{if } i=Z.\end{cases}$$

Now, the individuals' optimization problem can be represented as the familiar discrete choice problem:

$$\tag{3.14}V_{\boldsymbol{\alpha}}(i,\boldsymbol{\varepsilon}_{i,g})=\max_{a\in\mathcal{D}}\{\tilde{v}_{i,a}+\varepsilon_{i,a,g}\},$$

$$\tag{3.15}W_{\boldsymbol{\gamma}}(j,\boldsymbol{\varepsilon}_{j,h})=\max_{a\in\mathcal{D}}\{\tilde{w}_{a,j}+\varepsilon_{a,j,h}\},$$

where $\tilde{v}_{i,a}$ and $\tilde{w}_{a,j}$ are commonly referred to in the literature as the *choice-specific value functions*.

Let the conditional choice probability $\mathcal{P}_{i,j}$ denote the probability that choice j is the optimal choice for age i, that is, $\mathcal{P}_{i,j} = \int \mathbb{1}\{j = \arg\max_{a\in\mathcal{D}}(\tilde{v}_{i,a} + \varepsilon_{i,a,g})\}\mathfrak{f}(\boldsymbol{\varepsilon})\,d\boldsymbol{\varepsilon}$. Similarly, for females, $\mathcal{Q}_{i,j}$ is the probability that choice i is the optimal choice for age j females. That is, $\mathcal{Q}_{i,j} = \int \mathbb{1}\{i = \arg\max_{a\in\mathcal{D}}(\tilde{w}_{a,j} + \varepsilon_{a,j,h})\}\mathfrak{f}(d\boldsymbol{\varepsilon})$. The conditional choice probability can be expressed as a function of the normalized choice value functions, $(\tilde{v}_{i,j} - \tilde{v}_{i,0})$. In this case, the probability that a type i male who matches with a type j female will have the familiar multinomial logit form is

$$\mathcal{P}_{i,j} = \frac{\exp(\tilde{v}_{i,j} - \tilde{v}_{i,0})}{1 + \sum_{r=1}^{Z} \exp(\tilde{v}_{i,r} - \tilde{v}_{i,0})}. \tag{3.16}$$

Similarly, for females, the conditional choice probability is

$$\mathcal{Q}_{i,j} = \frac{\exp(\tilde{w}_{i,j} - \tilde{w}_{0,j})}{1 + \sum_{r=1}^{Z} \exp(\tilde{w}_{r,j} - \tilde{w}_{0,j})}. \tag{3.17}$$

I will now delve more deeply into the mean utilities $\tilde{v}_{i,j}$ and $\tilde{w}_{i,j}$ and derive a representation for the log-odds ratio of a match relative to remaining single. Consider the *integrated value function* as introduced by Rust (1987) where the unobserved state is integrated out of the Bellman equations. Let $\mathbf{V}_i$ and $\mathbf{W}_j$ be the integrated value function for a single age i male and age j female, respectively. That is, $\mathbf{V}_i = \mathbb{E}V_{\boldsymbol{\alpha}}(i, \boldsymbol{\varepsilon}_g) = \int V_{\boldsymbol{\alpha}}(i, \boldsymbol{\varepsilon}_g)\mathfrak{f}(\boldsymbol{\varepsilon}_g)\,d\boldsymbol{\varepsilon}_g$ and $\mathbf{W}_j = \mathbb{E}W_{\boldsymbol{\gamma}}(j, \boldsymbol{\varepsilon}_h) = \int W_{\boldsymbol{\gamma}}(j, \boldsymbol{\varepsilon}_h)\mathfrak{f}(\boldsymbol{\varepsilon}_h)\,d\boldsymbol{\varepsilon}_h$. In this finite-horizon case, the Type 1 Extreme Value distributional assumption allows these integrated value functions to have a recursive structure,

$$\mathbf{V}_i = \begin{cases} \alpha_{i,0} + c - \ln \mathcal{P}_{i,0} + \beta \mathbf{V}_{i+1}, & \text{if } i < Z, \\ \alpha_{i,0} + c - \ln \mathcal{P}_{i,0}, & \text{if } i = Z, \end{cases} \tag{3.18}$$

$$\mathbf{W}_j = \begin{cases} \gamma_{0,j} + c - \ln \mathcal{Q}_{0,j} + \beta \mathbf{W}_{j+1}, & \text{if } j < Z, \\ \gamma_{0,j} + c - \ln \mathcal{Q}_{0,j}, & \text{if } i = Z, \end{cases} \tag{3.19}$$

where c is Euler's constant. Appendix A.2 presents the derivations of Equations (3.18) and (3.19).

Consider the integrated value function for a male at terminal age Z. Equation (3.18) says that the expected value of participating in the marriage market at age Z is simply the familiar (McFadden's) social surplus function (or expected maximum utility) given by $\alpha_{Z,0} + c - \ln \mathcal{P}_{Z,0}$. It depends on the period utility from being single, $\alpha_{Z,0}$, and the equilibrium probability of a Z type male remaining single, $\mathcal{P}_{Z,0}$. If the marriage rate for age Z males is large (i.e., $\mathcal{P}_{Z,0}$ is small), then the expected value of participating in the marriage at age Z, $\mathbf{V}_Z$, will also be large. This familiar functional form arises from the Type 1 Extreme Value additively separable idiosyncratic payoff. Moving age backward to $i < Z$, the integrated value function in Equation (3.18) has an additional term represented by $\beta \mathbf{V}_{i+1}$. This is part of the mean utility from choosing to remain single at age i, and reflects the continuation value of participating in the marriage market in the next period as an older $(i+1)$ individual. When the individual is at a terminal age Z, the continuation value in the next period is hence zero. Equation (3.19) has an analogous interpretation.

For $i < Z$, after repeated substitution of Equation (3.18) into the choice-specific value function $\tilde{v}_{i,j}$ and $\tilde{v}_{i,0}$, and some algebra, the mean utilities can be expressed as a function of the expected surplus from the per-period random utilities:

$$\tilde{v}_{i,j} = \boldsymbol{\alpha}_i(j) - \tau_{i,j} + \sum_{k=1}^{T_{i,j}-1} \beta^k \big(1 - (1-\delta)^k\big)(\alpha_{i+k,0} + c - \ln \mathcal{P}_{i+k,0}) + \sum_{k=i+T_{i,j}}^{Z} \beta^{k-i}(\alpha_{k,0} + c - \ln \mathcal{P}_{k,0}) \mathbb{1}(j > i), \tag{3.20}$$

$$\tilde{v}_{i,0} = \alpha_{i,0} + \sum_{k=i+1}^{Z} \beta^{k-i}(\alpha_{k,0} + c - \ln \mathcal{P}_{k,0}). \tag{3.21}$$

Appendix A.3 presents the derivations of Equation (3.20). Similarly for $j < Z$, with repeated substitution of Equation (3.19) into the mean utilities from the female decision problem, $\tilde{w}_{i,j}$ and $\tilde{w}_{0,j}$, I get the following analogous expressions for the expected surplus to females:

$$\tilde{w}_{i,j} = \boldsymbol{\gamma}_j(i) + \tau_{i,j} + \sum_{k=1}^{T_{i,j}-1} \beta^k \big(1 - (1-\delta)^k\big)(\gamma_{0,j+k} + c - \ln \mathcal{Q}_{0,j+k}) + \sum_{k=j+T_{i,j}}^{Z} \beta^{k-j}(\gamma_{0,k} + c - \ln \mathcal{Q}_{0,k}) \mathbb{1}(j < i), \tag{3.22}$$

$$\tilde{w}_{0,j} = \gamma_{0,j} + \sum_{k=j+1}^{Z} \beta^{k-j}(\gamma_{0,k} + c - \ln \mathcal{Q}_{0,k}). \tag{3.23}$$

For $i = Z$, $\tilde{v}_{Z,j} = \boldsymbol{\alpha}_Z(j) - \tau_{Z,j}$ and $\tilde{v}_{Z,0} = \alpha_{Z,0}$, as defined in Equation (3.13). Similarly for $j = Z$, $\tilde{w}_{i,Z} = \boldsymbol{\gamma}_Z(i) + \tau_{i,Z}$ and $\tilde{w}_{0,Z} = \gamma_{0,Z}$, as defined in Equation (3.12).

The log-odds ratio of an (i, j) match relative to i remaining single identifies the difference in mean utilities, $\tilde{v}_{i,j} - \tilde{v}_{i,0}$. It describes the expected payoffs to an age i male from a match with an age j female relative to remaining single that period. This ratio has the following expression:

$$\log\left\{\frac{\mathcal{P}_{i,j}}{\mathcal{P}_{i,0}}\right\} = \begin{cases} \boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_{i,0}(j) - \tau_{i,j} \\ \quad - \sum_{k=0}^{T_{i,j}-1} \big(\beta(1-\delta)\big)^k \big(c + \ln \mathcal{P}_{i+k,0}^{-1}\big), & \text{if } \max(i,j) < Z, \\ \alpha_{i,j,1} - \alpha_{i,0} - \tau_{i,j}, & \text{if } \max(i,j) = Z. \end{cases} \tag{3.24}$$

Appendix A.4 presents the derivation of Equations (3.24) and (3.26). The log-odds ratio comprises three components: (1) the equilibrium transfer $\tau_{i,j}$, (2) the parameters

$$\boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_{i,0}(j) = \sum_{k=1}^{T_{i,j}} \big(\beta(1-\delta)\big)^{k-1} (\alpha_{i,j,k} - \alpha_{i+k-1,0}), \tag{3.25}$$

representing the discounted stream of within-marriage net utilities an age i male gets in the kth period of marriage relative to the per-period utility from remaining single that period (in the event that the marriage does not dissolve), and (3) the term $-\sum_{k=1}^{T_{i,j}-1} (\beta(1-\delta))^k (c + \ln \mathcal{P}_{i+k,0}^{-1})$, representing the discounted sum of future log male marriage rates.[22] Higher future marriage rates increase the opportunity cost of committing to an (i, j) match today, thus lowering the mean utility of the match relative to remaining single.[23]

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 $\tilde{w}_{i,j} - \tilde{w}_{i,0}$. That is:

$$\log\left\{\frac{\mathcal{Q}_{i,j}}{\mathcal{Q}_{0,j}}\right\} = \begin{cases} \boldsymbol{\gamma}_j(i) - \boldsymbol{\gamma}_{0,j}(i) + \tau_{i,j} \\ \quad - \sum_{k=1}^{T_{i,j}-1} \big(\beta(1-\delta)\big)^k \big(c + \ln \mathcal{Q}_{0,j+k}^{-1}\big), & \text{if } \max(i,j) < Z, \\ \boldsymbol{\gamma}_j(i) - \boldsymbol{\gamma}_{0,j}(i) + \tau_{i,j}, & \text{if } \max(i,j) = Z. \end{cases} \tag{3.26}$$

[22]When the marriage rates at $i + k$ are high, so would be $\ln \mathcal{P}_{i+k,0}^{-1}$.

[23]Notice that the kth period discount weights in the log-odds ratio, $(\beta(1-\delta))^{k-1}$, reflect time discounting of future period utilities (β^{k-1}), as well as the probability that the marriage survives to that period given by $(1-\delta)^{k-1}$.

Equation (3.26) gives the difference in systematic expected payoffs for a j type female marrying an i type male relative to remaining single during that same period. The term

$$\gamma_j(i) - \gamma_{0,j}(i) = \sum_{k=1}^{T_{i,j}} \big(\beta(1-\delta)\big)^{k-1}(\gamma_{i,j,k} - \gamma_{0,j+k-1}) \tag{3.27}$$

represents the discounted stream of within-marriage net utilities accrued to an age j female in the kth period of marriage relative to the per-period utility from remaining single during that same period. Contrary to the male's equation, the transfer $\tau_{i,j}$ enters additively. Just as in the male case, the term $\sum_{k=1}^{T_{i,j}-1}(\beta(1-\delta))^k(c + \ln \mathcal{Q}_{0,j+k}^{-1})$ captures the opportunity cost incurred from choosing to marry at age j instead of staying single and participating in the marriage market in the future.

3.3. *Equilibrium and the Dynamic Marriage Matching Function*

Rearranging the terms of Equations (3.24) and (3.26) delivers a dynamic analogue of the log-odds ratio. The left-hand side of Equation (3.28) has the natural log of the probability that an age i male matches with an age j female, $\mathcal{P}_{i,j}$, relative to the scaled products of the probabilities that an age i male remains single during that period and for the duration of the proposed match, $\prod_{k=0}^{T_{i,j}-1} \mathcal{P}_{i+k,0}^{(\beta(1-\delta))^k}$. 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. The future probabilities of remaining single at age $i+k$ are scaled by the discount factor and the probabilities that the marriage survives for k periods. I will refer to this quantity as the *dynamic log-odds ratio* for an age i male marrying with an age j female. The constant κ is the geometric sum of Euler's constants, $\kappa = c\beta(1-\delta)(1-(\beta(1-\delta))^{T_{i,j}})/(1-\beta(1-\delta))$:

$$\ln \mathcal{P}_{i,j} - \sum_{k=0}^{T_{i,j}-1} \big(\beta(1-\delta)\big)^k \ln \mathcal{P}_{i+k,0} = \boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_{i,0}(j) - \tau_{i,j} - \kappa, \tag{3.28}$$

$$\ln \mathcal{Q}_{i,j} - \sum_{k=0}^{T_{i,j}-1} \big(\beta(1-\delta)\big)^k \ln \mathcal{Q}_{0,j+k} = \boldsymbol{\gamma}_{0,j}(i) - \boldsymbol{\gamma}_{0,j}(i) + \tau_{i,j} - \kappa. \tag{3.29}$$

Equations (3.28) and (3.29) identify the discounted stream of within-marriage net utilities an age i male gets relative to the per-period utility from remaining single, $\boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_{i,0}(j)$, less the equilibrium transfer, $\tau_{i,j}$. The interpretation of the dynamic log-odds ratio for an age j female marrying an age i male in Equation (3.29) is similar, with the central difference being that the age j female is the recipient of the transfer. I will denote the dynamic log-odds for an age i male marrying an age j female by $n_{i,j}(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f}) = \ln(\mathcal{P}_{i,j}/\prod_{k=0}^{T_{i,j}-1} \mathcal{P}_{i+k,0}^{(\beta(1-\delta))^k})$, and for the female counterpart, $\mathcal{N}_{i,j}(\boldsymbol{\mu}, \mathbf{m}, \mathbf{f}) = \ln(\mathcal{Q}_{i,j}/\prod_{k=0}^{T_{i,j}-1} \mathcal{Q}_{0,j+k}^{(\beta S)^k})$.

Since Equations (3.28) and (3.29) are defined for every (i, j) marriage pair, in aggregate they form a system of $(Z \times Z)$ quasi-demand and quasi-supply equations, respectively.[24] I now define the Walrasian marriage market equilibrium.

DEFINITION 1: A *marriage market equilibrium* consists of a vector of available males, $\mathbf{m}$, and females, $\mathbf{f}$, the vector of marriages, $\boldsymbol{\mu}$, and the vector of transfers, $\boldsymbol{\tau}$, such that the number of age i men who want to marry age j spouses exactly equals the number of age j women who agree to marry age i men for all combinations of (i, j). That is, for each of the $(Z \times Z)$ sub-markets,

$$m_i\mathcal{P}_{i,j} = f_j\mathcal{Q}_{i,j} = \mu_{i,j}.$$

Dynamic Marriage Matching Function

Let $p_{i,j}$ and $q_{i,j}$ denote the maximum likelihood (ML) estimators of the probability that an age i male matches with an age j female, $\mathcal{P}_{i,j}$, and an age j female matches with an age i male, $\mathcal{Q}_{i,j}$, respectively. That is, $p_{i,j} = \mu_{i,j}/m_i$ and $q_{i,j} = \mu_{i,j}/f_j$. The above marriage market clearing conditions and the ML estimators for the choice probabilities are applied to the system of quasi-supply and demand equations, (3.28) and (3.29) respectively, to derive the Dynamic Marriage Matching Function for an (i, j) marriage (when $T_{i,j} > 0$)[25] given by Equation (3.1):

$$\mu_{i,j} = \widehat{\Pi}_{i,j}\sqrt{m_i f_j}\prod_{k=0}^{T_{i,j}-1}\left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k}f_{j+k}}\right)^{(1/2)(\beta(1-\delta))^k},$$

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

[24]In the static framework of CS, we get an analogous representation of quasi-demand and quasi-supply of spouses corresponding to the case when $T_{i,j} = 0$. That is:

$$\ln \mathcal{P}_{i,j} - \ln \mathcal{P}_{i,0} = \alpha_{i,j} - \alpha_{i,0} - \tau_{i,j},$$
$$\ln \mathcal{Q}_{i,j} - \ln \mathcal{Q}_{0,j} = \gamma_{i,j} - \gamma_{0,j} + \tau_{i,j}.$$

[25]This condition ensures that neither spouse is at a terminal age. If $T_{i,j} = 0$, the Dynamic Marriage Matching Function reduces to the static marriage matching function of CS, which is $\mu_{i,j} = \widehat{\Pi}_{i,j}\sqrt{\mu_{i,0}\mu_{0,j}}$.

The dynamic marriage matching function also needs to satisfy the accounting constraints given by Equations (3.30), (3.31), and (3.32):

$$\mu_{0,j} + \sum_{i=1}^{Z} \mu_{i,j} = f_j \quad \forall j, \tag{3.30}$$

$$\mu_{i,0} + \sum_{j=1}^{Z} \mu_{i,j} = m_i \quad \forall i, \tag{3.31}$$

$$\mu_{0,j}, \mu_{i,0}, \mu_{i,j} \geq 0 \quad \forall i, j. \tag{3.32}$$

Equation (3.30) states that the total number of age j women who marry and the number of unmarried age j women must be equal to the number of available age j women for all j. Similarly, Equation (3.31) states that the total number of women who marry age i men and the number of unmarried age i men must be equal to the number of available age i men for all i. Equation (3.32) holds because the number of unmarrieds of any age and gender and the number of marriages between age i men and age j women must be nonnegative.

Azevedo and Hatfield (2013) demonstrated that, in a quasilinear setting, a competitive equilibrium exists in two-sided markets with a continuum of agents with arbitrary preferences. In large two-sided markets with a finite number of agents of each type and quasilinear utility like that of this paper's model, the authors showed that equilibria exist that approximately clear the market. I refer readers to the results and proof in that paper.

Given the preference parameters of the system, $\boldsymbol{\Pi}$, practitioners are often interested in how variations in the supply population vectors, $\mathbf{m}$ and $\mathbf{f}$, affect the distribution of marriages as represented by $\boldsymbol{\mu}$. I refer to this as the Dynamic Marriage Matching (DMM) 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 nonnegative, and strictly positive population vectors, $\mathbf{m}$ and $\mathbf{f}$, there exists a unique nonnegative marital distribution $\boldsymbol{\mu}$ that is consistent with $\boldsymbol{\Pi}$, and that satisfies Equations (3.30), (3.31), (3.32), and (3.26).

Taking $\widehat{\Pi}_{i,j}$, $\mathbf{m}$, and $\mathbf{f}$ as exogenously given, Equation (3.1) defines a $Z \times Z$ system of polynomials with the $Z \times Z$ elements of $\boldsymbol{\mu}$ as unknowns. The exogeneity of the number of single individuals, $\mathbf{m}$ and $\mathbf{f}$, allows us to reformulate the model as a $2Z$ system of polynomials with $2Z$ number of unmarrieds by age, $\mu_{i,0}$ and $\mu_{0,j}$, as unknowns. I derive this "reduced" system as defined by Equations (3.33) and (3.34) below by summing Equation (3.1) over all i's and Equation (3.1) over all j's, respectively. The $2Z$ system represented by Equations (3.33) and (3.34) has a significant computational advantage when solving the inverse problem of calculating $\mu_{i,0}$ and $\mu_{0,j}$, consistent with an estimate of $\boldsymbol{\Pi}$ and a vector of single individuals, $\mathbf{m}$ and $\mathbf{f}$. Equation (3.1) fully determines the distribution of marriage $\boldsymbol{\mu}$ after solving for $\mu_{i,0}$ and $\mu_{0,j}$:

$$m_i - \mu_{i,0} = \sum_{i=1}^{I} \widehat{\Pi}_{i,j} \sqrt{m_i f_j} \prod_{k=0}^{T_{i,j}-1} \left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k} f_{j+k}} \right)^{(1/2)(\beta(1-\delta))^k}, \tag{3.33}$$

$$f_j - \mu_{0,j} = \sum_{j=1}^{J} \widehat{\Pi}_{i,j} \sqrt{m_i f_j} \prod_{k=0}^{T_{i,j}-1} \left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k} f_{j+k}} \right)^{(1/2)(\beta(1-\delta))^k}. \tag{3.34}$$

3.4. *Identification*

3.4.1. *Dynamic Marriage Gains*

CS 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:

$$\ln \frac{\mu_{i,j}^2}{\mu_{i,0}\mu_{0,j}} = \ln \frac{p_{i,j} q_{i,j}}{p_{i,0} q_{0,j}} = (\alpha_{i,j} + \gamma_{i,j}) - (\alpha_{i,0} + \gamma_{0,j}) = 2 \ln \widehat{\pi}_{i,j}. \tag{3.35}$$

The dynamic analogue of this statistic derived from Equation (3.1) takes the form

$$\begin{aligned} &\ln \left(p_{i,j} q_{i,j} \Big/ \prod_{k=0}^{T_{i,j}-1} (p_{i+k,0} q_{0,j+k})^{(\beta(1-\delta))^k} \right) \\ &\quad = \boldsymbol{\alpha}_i(j) + \boldsymbol{\gamma}_j(i) - \boldsymbol{\alpha}_i(0) - \boldsymbol{\gamma}_j(0) - 2\boldsymbol{\kappa} = 2 \ln \widehat{\Pi}_{i,j}. \end{aligned} \tag{3.36}$$

The interpretation of the statistic is similar to the static case. $2 \ln \widehat{\Pi}_{i,j}$ gives the couple's present discounted utility from being locked in an (i, j) match today for the duration of the match relative to the present discounted sum of the per-period payoff from being single for that same time span. This statistic is point identified from a single cross-section of data on aggregate marriages. The right-hand side of Equation (3.36) is composed of the primitives of the model only, and these are invariant to changes in the vectors of unmarried men, $\mathbf{m}$, and women, $\mathbf{f}$. This statistic becomes the basis of the empirical application in Section 5. I will refer to the statistic $2 \ln \pi_{i,j}$ as defined by Equation (3.35) as *Static Gains* and to $2 \ln \Pi_{i,j}$ from Equation (3.36) as *Dynamic Gains*.

3.4.2. *A Test of the Model*

Equations (3.28) and (3.29) can be expressed in terms of the ML estimators $p_{i,j}$ and $q_{i,j}$. That is,

$$\ln\left(p_{i,j}\Big/\prod_{k=0}^{T_{i,j}-1} p_{i+k,0}^{(\beta(1-\delta))^k}\right)=\boldsymbol{\alpha}_i(j)-\boldsymbol{\alpha}_i(0)-\tau_{i,j}-\kappa, \tag{3.37}$$

$$\ln\left(q_{i,j}\Big/\prod_{k=0}^{T_{i,j}-1} q_{0,j+k}^{(\beta(1-\delta))^k}\right)=\boldsymbol{\gamma}_j(i)-\boldsymbol{\gamma}_j(0)+\tau_{i,j}-\kappa. \tag{3.38}$$

Let

$$n_{i,j}(\boldsymbol{\mu},\mathbf{m},\mathbf{f})=\ln\left(p_{i,j}\Big/\prod_{k=0}^{T_{i,j}-1} p_{i+k,0}^{(\beta(1-\delta))^k}\right)\quad\text{and}$$

$$\mathcal{N}_{i,j}(\boldsymbol{\mu},\mathbf{m},\mathbf{f})=\ln\left(q_{i,j}\Big/\prod_{k=0}^{T_{i,j}-1} q_{0,j+k}^{(\beta(1-\delta))^k}\right).$$

Proposition 1 below provides a simple test for our model:

PROPOSITION 1: *Holding $\alpha_{i,jk}$, $\gamma_{i,jk}$, and δ fixed for all (i,j,k), any changes in available men m_i or women f_j leading to a non-decreasing change in $n_{i,j}(\boldsymbol{\mu},\mathbf{m},\mathbf{f})$ would also lead to a non-increasing change in $\mathcal{N}_{i,j}(\boldsymbol{\mu},\mathbf{m},\mathbf{f})$, and vice versa.*

In other words, any changes in the relative scarcity of men and women that affect the market clearing division of surplus $\tau_{i,j}$ would make $n_{i,j}$ and $\mathcal{N}_{i,j}$ move in opposite directions. If our model is true, a simple regression of estimates of $\hat{n}_{i,j}(\boldsymbol{\mu},\mathbf{m},\mathbf{f})$ against $\hat{\mathcal{N}}_{i,j}(\boldsymbol{\mu},\mathbf{m},\mathbf{f})$ should yield a slope coefficient of -1.[26]

4. SUPPLIES, COMMITMENT, AND DIVORCE

4.1. *Endogenizing Supplies*

The model so far assumes that the vector of single available individuals at the beginning of each period is given by **m** and **f**. This assumption is convenient for several reasons. For one, when faced with data on the number of single individuals by age, it allows me to abstract from the economic and demographic

[26]Changes in m_i or f_j that leave $\tau_{i,j}$ unchanged would also not affect $n_{i,j}(\boldsymbol{\mu},\mathbf{m},\mathbf{f})$ and $\mathcal{N}_{i,j}(\boldsymbol{\mu},\mathbf{m},\mathbf{f})$.

factors that relate $(m_{i'}, f_{j'})$ to (m_i, f_j), where $i \neq i$ and $j \neq j'$, and ignore mortality and morbidity factors that become important as an individual ages. Data also present prominent inflow and outflow migrations at younger ages. Computationally, the assumption also allows me to reduce the model's dimensionality by expressing marriage matching functions in terms of the number of unmarried individuals, as described in the previous section.

A more realistic assumption would be to allow the number of single individuals by age to be endogenously determined by marriage decisions. Consider a similar environment to that described in Section 3 where only m_1 and f_1 individuals are exogenous. That is, these are the number of year 1 individuals born into the marriage market. With a little algebra, I can derive the equations of motion for the number of single available men and women:

$$m_{k+1} = \begin{cases} m_k - \sum_{j=1}^{Z-1} (1-\delta)\mu_{k,j}, & \text{if } k = 1, \\ m_k - \sum_{j=1}^{Z-1} (1-\delta)\mu_{k,j} + \sum_{i=1}^{k-1} \sum_{l=1}^{Z-1} (1-\delta)^{k-i} \delta \mu_{i,l}, & \text{if } k > 1, \end{cases} \tag{4.1}$$

$$f_{k+1} = \begin{cases} f_k - \sum_{i=1}^{Z-1} (1-\delta)\mu_{i,k}, & \text{if } k = 1, \\ f_k - \sum_{i=1}^{Z-1} (1-\delta)\mu_{i,k} + \sum_{j=1}^{k-1} \sum_{l=1}^{Z-1} (1-\delta)^{k-j} \delta \mu_{l,j}, & \text{if } k > 1. \end{cases} \tag{4.2}$$

These equations are mirror images of one another. The second term accounts for the number of individuals who got married in the previous period and are still married. These individuals leave the pool of single individuals participating in the marriage market during this period at age $k+1$. The third term accounts for individuals who return to the marriage market as aged $k+1$ individuals after the dissolution of earlier marriages.

The model as outlined in Section 3 assumes a society of overlapping generations of adults whose decisions generate a stationary distribution of new marriages. In the case where the number of single individuals is endogenously determined, this flow of new marriages is sustained by the exogenously given number of m_1 and f_1 individuals born to the society in each period. A number of feasibility conditions need to hold to ensure the existence of a stationary distribution of marriages. On top of the accounting constraints given by Equations (3.30), (3.31), and (3.32), tuple $(m_1, f_1, \delta, \boldsymbol{\mu}, \boldsymbol{\Pi})$ must be such that Equations (4.1) and (4.2) generate $m_k \geq 0$, and $f_k \geq 0$, for all $k > 1$. After some tedious algebra and repeated substitution of Equations (4.1) and (4.2), I arrive at the following dynamic constraints for $k = 2, \dots, Z$:

$$m_1 \geq \sum_{j=1}^{Z-1} (1-\delta)\mu_{k-1,j} + \sum_{i=2}^{k-1} \sum_{j=1}^{Z-i} (1-\delta)^i \mu_{k-i,j}, \tag{4.3}$$

$$f_1 \geq \sum_{i=1}^{Z-1} (1-\delta)\mu_{i,k-1} + \sum_{j=2}^{k-1} \sum_{i=1}^{Z-j} (1-\delta)^j \mu_{i,k-j}. \tag{4.4}$$

The right-hand side of these constraints gives the number of age k individuals who got married in the previous periods and are still married. These constraints say that the number of age k individuals who are still married cannot exceed the initial m_1 and f_1 born to the system.

In this case, where supplies are endogenous, the marriage matching function in Equation (3.1) needs to be augmented with Equations (4.1) and (4.2) to complete the model. The model now comprises a system of $Z \times Z$ polynomial equations in $\boldsymbol{\mu}$. The estimation of $\Pi_{i,j}$ would require that I first compute the vector of available men and women by age, $\mathbf{m}$ and $\mathbf{f}$. Hence, given data on $(\boldsymbol{\mu}, m_1, f_1)$, I would first need to compute the vector of $\mathbf{m}$ and $\mathbf{f}$, as implied by these demographic equations of motion, before computing the marriage gains. The empirical application in Section 5 allows for endogenous supplies together with differential mortality and heterogeneity in the divorce rates.

4.2. *Commitment and Divorce—Generalizing the Divorce Hazard*

The model assumes full commitment where couples commit to an allocation of marital surplus that is predetermined by their respective ages at marriage. Divorce is an exogenous shock that dissolves otherwise fully committed stable couples. This simplification is integral to identifying the net gains and the total marriage gains from a single cross-section of data on aggregate marriage behaviors.

The literature has long recognized that marital dissolution and the level of marital commitment are clearly interconnected. Life-cycle decisions on labor supply, savings, and human-capital accumulations affect individuals' outside options and accordingly their relative marital bargaining power. The full-commitment assumption allows me to side-step these important though complicating considerations. A cost of this approach is that the model has nothing to say about the division of within-marriage surplus over the life-cycle, or how that is affected by changes in the labor market or the marriage market. The identification strategy is predicated on the observation that the aggregate number of new marriages, suitably scaled by the proportion of individuals who stay single at their current and future ages, identifies the marriage surplus parameter. Any changes in the marriage market that affect the outside option do so through the number of available individuals in the marriage market. A more complete equilibrium model that allows for limited commitment within marriage would, of course, require more detailed labor supply data for different couple types. The model here is therefore a building-block toward such an analysis.[27]

So far, the model has assumed that divorce is exogenous and occurs at a constant rate of δ. The heterogeneity in divorce patterns in the United States has been well studied. For example, using data from the National Longitudinal Study of the High School Class of 1972, Weiss and Willis (1997) identified a number of interesting patterns. The authors found that the divorce hazard initially increases with the duration of a match before decreasing. Couples with similar education levels, religion, and ethnicity at the time of marriage are less likely to divorce. While allowing for the endogeneity of divorce is beyond this paper's scope, I can allow the divorce hazard to reflect some of the heterogeneity observed in data. One way would be to allow the divorce hazard to depend on the ages of the couple and duration of the match.

For a couple who marry at age (i, j), let $\delta_{i,j,d}$ be the exogenous divorce hazard in the dth year of the marriage where $1 \leq d \leq T_{i,j}$. Hence, this marriage will survive until year d with an unconditional probability given by $\prod_{l=1}^{d}(1-\delta_{i,j,l})$. Divorce couples re-enter the marriage market without penalty.[28] Working through the algebra of the model, the corresponding log-odds that allow for heterogeneity in the divorce hazard (analogous to Equations (3.24) and (3.26)) are

$$\log\left\{\frac{\mathcal{P}_{i,j}}{\mathcal{P}_{i,0}}\right\} = \begin{cases} \boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_{i,0}(j) - \tau_{i,j} \\ \quad - \displaystyle\sum_{k=1}^{T_{i,j}-1} \beta^k \prod_{l=1}^{k}(1-\delta_{i,j,l})\left(c + \ln \mathcal{P}_{i+k,0}^{-1}\right), & \text{if } \max(i,j) < Z, \\ \boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_{i,0}(j) - \tau_{i,j}, & \text{if } \max(i,j) = Z, \end{cases} \tag{4.5}$$

[27]A number of papers allow for intra-household allocation and the level of commitment to be jointly determined. Using data from the Panel Study of Income Dynamics, Mazzocco, Ruiz, and Yamaguchi (2007) estimated a dynamic model to examine how labor supply and wealth accumulation decisions change with transition in and out of marriage. In this model, married couples cooperate in their decision making without committing to future allocations of marital surplus. Divorce happens when no within-marriage reallocation can make the married couple better off than when single. Iyigun (2009) proposed a theoretical model of intra-household allocation where couples decide on the level of commitment and whether to marry or cohabitate. The paper investigates how individuals sort into marriage and cohabitation with different commitment levels as the costs of commitment and marital preferences vary across individuals.

[28]This simplification can be relaxed. Assuming it is possible to differentiate the "previously married" from the "never married" individuals, the state vector can be extended to allow for this status. Accordingly, I can also allow for the parameters to be indexed by this additional state.

for an aged i male marrying an aged j female relative to remaining single. And for an aged j female marrying an aged i male relative to remaining single, the log-odds are

$$\log\left\{\frac{\mathcal{Q}_{i,j}}{\mathcal{Q}_{0,j}}\right\} = \begin{cases} \boldsymbol{\gamma}_j(i) - \boldsymbol{\gamma}_{0,j}(i) + \tau_{i,j} \\ \quad - \sum_{k=1}^{T_{i,j}-1} \beta^k \prod_{l=1}^{k} (1-\delta_{i,j,l})\big(c + \ln \mathcal{Q}_{0,j+k}^{-1}\big), & \text{if } \max(i,j) < Z, \\ \boldsymbol{\gamma}_j(i) - \boldsymbol{\gamma}_{0,j}(i) + \tau_{i,j}, & \text{if } \max(i,j) = Z. \end{cases} \tag{4.6}$$

Solving for the equilibrium number of marriages, I obtain an alternate marriage matching function in Equation (4.2). If I continue to assume $\mathbf{m}$ and $\mathbf{f}$ are exogenous, I can still employ the dimensionality reducing transformation where the system of equations is expressed in terms of the equilibrium number of unmarried individuals such as in Equations (3.33) and (3.34):

$$\mu_{i,j} = \Pi_{i,j}\sqrt{m_i f_j} \prod_{k=0}^{T_{i,j}-1} \left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k} f_{j+k}}\right)^{(1/2)\beta^k \prod_{l=1}^{k}(1-\delta_{i,j,l})},$$

where $\ln \Pi_{i,j} = \frac{1}{2}(\boldsymbol{\alpha}_i(j) + \boldsymbol{\gamma}_j(i) - \boldsymbol{\alpha}_{i,0}(j) - \boldsymbol{\gamma}_{0,j}(i)) - \kappa$.[29]

5. EMPIRICAL APPLICATION

In this empirical application of the model, I relax a number of assumptions made in Section 3 for expositional convenience. I allow for

1. differential mortality by age and gender,
2. divorce rate heterogeneity by age of couple and duration of the match,
3. and endogenous supply of available men and women.

Consider a system where only the number of men and women, m_1 and f_1, born each period is exogenous. Individuals are assumed to live for $Z = 60$ periods. At each period, let the mortality rate for an age i male be $(1 - \varrho_i)$; that is, the probability that an age i male is alive at the beginning of the next period is ϱ_i. For an age j female, I denote the mortality rate as $(1 - \rho_j)$. The mortality rates are taken from the Life Tables of the "Vital Statistics of the United States" published by the U.S. Department of Health and Human Services. Since the model is used to analyze the changes in the U.S. marriage

[29]Details on these present discounted within-marriage utilities are provided in footnote 36. In this case where there is no differential mortality, the age-specific survival probabilities, ρ_l and ϱ_l, have been set to 1.

distribution over two decades from 1970 to 1990, the mortality rates are taken from the Life Tables for 1970, 1980, and 1990.[30]

For a couple who marry at age (i, j), let $\delta_{i,j,d}$ be the exogenous divorce hazard in the d'th year of the marriage where $1 \leq d \leq T_{i,j}$. These divorce rates are estimated using data from divorce records from 1968 to 1995, in the Divorce Public-Use Tape Files published by the National Center of Health Statistics.[31] A simple logit regression with age polynomials and duration fixed effect is fitted over the divorce data. The estimated model is then used to get a predicted estimate of $\delta_{i,j,d}$. Since there have been well-documented changes in divorce rates over the period of 1970 to 1990, I also allow for time fixed effect. The details and results of the divorce rate estimation are presented in the Supplemental Material.

For completeness, I state the equations used in my empirical application. These equations mirror their counterparts described in Section 3. The dynamic marriage matching function allowing for differential mortality and divorce rate heterogeneity takes the form

$$\mu_{i,j} = \Pi_{i,j}\sqrt{m_i f_j} \prod_{k=0}^{T_{i,j}-1} \left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k} f_{j+k}}\right)^{(1/2)\beta^k \prod_{l=1}^{k} \varrho_l \rho_l (1-\delta_{i,j,k})}, \tag{5.1}$$

where $\ln \Pi_{i,j} = \frac{1}{2}(\boldsymbol{\alpha}_i(j) + \boldsymbol{\gamma}_j(i) - \boldsymbol{\alpha}_{i,0}(j) - \boldsymbol{\gamma}_{0,j}(i)) - \kappa$.[32]

[30]While I allow for variation of mortality rates by gender and time, I ignore state variation in mortality rates.

[31]These data are available from the data link on the National Bureau of Economic Research web-site (http://www.nber.org/data/marrdivo.html).

[32]The present discounted utility within marriage with differential mortality rates and divorce rate heterogeneity are

$$\boldsymbol{\alpha}_i(j) = \alpha_{i,j,1} + \sum_{k=1}^{T_{i,j}-1} \beta^k \prod_{l=1}^{k} \varrho_l \rho_l (1-\delta_{i,j,l}) \alpha_{i,j,k+1},$$

$$\boldsymbol{\alpha}_{i,0}(j) = \alpha_{i,0} + \sum_{k=1}^{T_{i,j}} \beta^k \prod_{l=1}^{k} \varrho_l \rho_l (1-\delta_{i,j,l}) \alpha_{i+k,0},$$

$$\boldsymbol{\gamma}_j(i) = \gamma_{i,j,1} + \sum_{k=1}^{T_{i,j}-1} \beta^k \prod_{l=1}^{k} \varrho_l \rho_l (1-\delta_{i,j,l}) \gamma_{i,j,k+1},$$

$$\boldsymbol{\gamma}_{0,j}(i) = \gamma_{0,j} + \sum_{k=1}^{T_{i,j}} \beta^k \prod_{l=1}^{k} \varrho_l \rho_l (1-\delta_{i,j,l}) \gamma_{0,j+k}.$$

The equations determining the number of available single men and women for ages $2 \leq k \leq Z$ at each period are given by

$$m_{k+1} = \varrho_k \Bigg[m_k - \sum_{j=1}^{Z-1} \rho_j (1 - \delta_{k,j,1}) \mu_{k,j} + \sum_{i=1}^{k-1} \sum_{l=1}^{Z-1} \big(1 - \rho_{l+k-i}(1 - \delta_{i,l,k+i+1})\big) \times \left\{ \prod_{r=0}^{k-i-1} \varrho_{i+r} \rho_{l+r} (1 - \delta_{i,l,r+1}) \right\} \mu_{i,l} \Bigg], \tag{5.2}$$

$$f_{k+1} = \rho_k \Bigg[f_k - \sum_{i=1}^{Z-1} \varrho_i (1 - \delta_{i,k,1}) \mu_{i,k} + \sum_{j=1}^{k-1} \sum_{l=1}^{Z-1} \big(1 - \varrho_{l+k-j}(1 - \delta_{l,j,k+j+1})\big) \times \left\{ \prod_{r=0}^{k-j-1} \rho_{j+r} \varrho_{l+r} (1 - \delta_{l,j,r+1}) \right\} \mu_{l,j} \Bigg]. \tag{5.3}$$

The equations closely mirror those discussed on Section 4.1. The accompanying dynamic constraints are

$$\prod_{g=1}^{k-2} \varrho_g m_1 \geq \sum_{j=1}^{Z-1} \rho_j (1 - \delta_{k-1,j,1}) \mu_{k-1,j} + \sum_{i=2}^{k-1} \sum_{j=1}^{Z-i} \prod_{h=0}^{i-2} \varrho_{k-2-h} \prod_{l=0}^{i-1} \rho_{j+l} (1 - \delta_{k-i,j,1+l}) \mu_{k-i,j}, \tag{5.4}$$

$$\prod_{g=1}^{k-2} \rho_g f_1 \geq \sum_{i=1}^{Z-1} \varrho_i (1 - \delta_{i,k-1,1}) \mu_{i,k-1} + \sum_{j=2}^{k-1} \sum_{i=1}^{Z-j} \prod_{h=0}^{j-2} \rho_{k-2-h} \prod_{l=0}^{j-1} \varrho_{i+l} (1 - \delta_{i,k-j,1+l}) \mu_{i,k-j}, \tag{5.5}$$

where $k = 2, \ldots, Z$. Both m_1 and f_1 need to satisfy the dynamic constraints (5.4) and (5.5), and the accounting flow Equations (5.2) and (5.3), which endogenously determine the number of available men, m_k, and women, f_k, respectively. The levels of m_1 and f_1 (i.e., the number of available men and women at 16 years of age) taken directly from the U.S. Census, however, fail the dynamic constraints generating level of available men and women that could not sustain the flow of new marriages, $\boldsymbol{\mu}$. Instead, m_1 and f_1 are estimated as the lowest number of individuals entering the system that satisfy both the accounting and dynamic constraints. Details of the estimation are provided in Appendix A.5.

To identify the vector of marital gains, $\widehat{\boldsymbol{\Pi}}$, the minimum levels of m_1 and f_1 satisfying the dynamic and accounting constraints are first estimated. Using Equations (5.2) and (5.3), I then identify the stationary level of m_k and f_k for ages $k = 2, \ldots, Z$ and the corresponding levels of $\{\mu_{i,0}, \mu_{0,j}\}$ implied by the observed distribution of new marriages $\boldsymbol{\mu}$. Using Equation (5.1), I can express $\Pi_{i,j}$ as a function of observed quantities providing an estimate of the marital gains for an (i, j) marriage.

5.1. *Changes in the U.S. Marriage Distribution, 1970–1990*

The model is used to analyze the changes in the U.S. marriage distribution 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 marriage gains. I construct the distribution of new marriages and single available men and women aged 16–75 over the two decades for the United States. To minimize sparseness in the marriage distribution, a two-year distribution of new marriages is used (instead of a one-year). The marriage distributions by age $\widehat{\mu}_{i,j}$ for 1971/72, 1981/82, and 1991/92 are constructed using the Vital Statistics data taken from the N.B.E.R. collection of the National Center for Health Statistics.[33] These files contain a sample of the new marriage records from 42 reporting states across the three periods.[34] The numbers of single available men and women by age, $\widehat{\mu}_{i,0}$ and $\widehat{\mu}_{0,j}$, come from the 1970, 1980, and 1990 U.S. Census. To be consistent, only unmarried individuals from matching reporting states are included.[35] The individuals are aged between 16 and 75. An individual is considered unmarried if his or her marital status is (i) not equal to married spouse

[33]These data are collected by the U.S. Department of Health and Human Services.

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

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

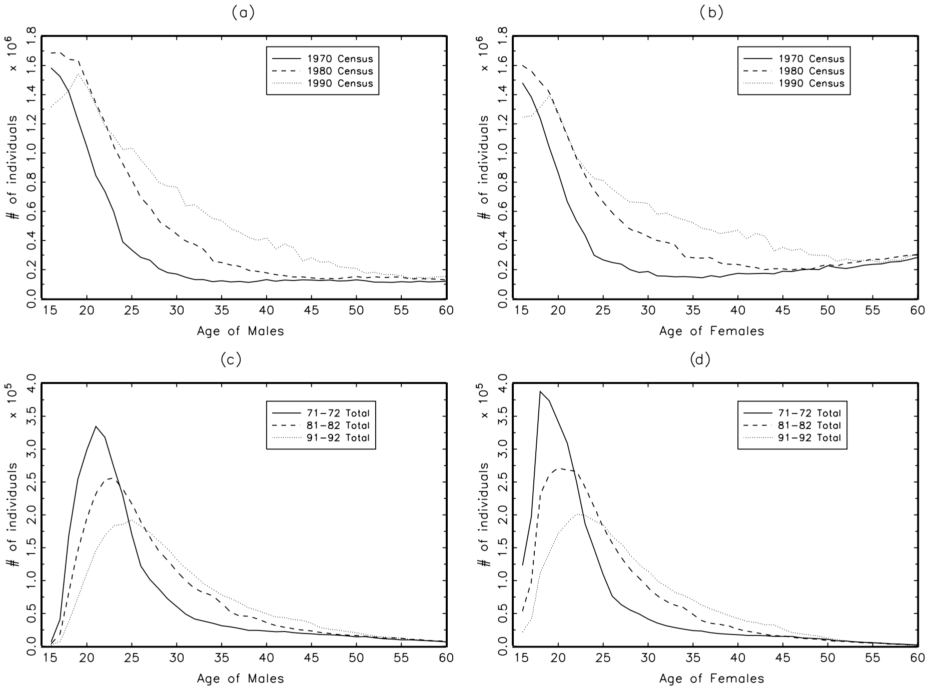

FIGURE 2.—(a) Single males from Census data. (b) Single females from Census data. (c) Married males from Vital Statistics. (d) Married females from Vital Statistics.

present, or (ii) married spouse absent. Census weights are used to get an estimate of the total unmarried counts.

Figures 2(a) and 2(b) plot the distribution of single men and women according to the three U.S. Censuses. The plots show some familiar patterns. There are fewer single individuals at older ages than at younger ages. Moving from 1970 to 1980 and 1990, there is a dramatic change in the number of available single men and women as the baby boomers enter marriageable age. Gender differences in later ages arise due to higher mortality rates among older men and lower remarriage rates among divorced women. Figures 2(c) and 2(d) graph the marginal distribution of two-year new marriages over the period. There is a clear shift to the right in the distribution of new marriages across genders due to more delayed marriage. The modal age of marriage across gender also increases. For males, the modal age of marriage goes from approximately 21 in 1971/72 to 23 in 1981/82 and to 25 in 1991/92. The modal age for females, which is slightly younger compared to males, also increases from approximately 18 in 1971/72 to 20 in 1981/82 and to 22 in 1991/92.

Table I provides summary statistics of the data. According to the 1970 Census, there were 16.0 million and 19.6 million single men and women, respectively, between the ages of 16 and 75. By 1980, the number of available men

TABLE I

DATA SUMMARY

A: U.S. 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

	1971/72	1981/82	1991/92
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

and women had increased by 46.2% and 39%, respectively, to 23.4 million men and 27.2 million women. 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 were 3.24 million new marriages recorded in the two years 1971/72, while in 1981/82 there were 3.45 million new marriages. This is an increase of 6.5% compared to the approximately 40% increase in the 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.

5.2. *Estimating Marriage Gains*

Figure 3(a) graphs the distribution of new marriages by age in 1971/72, $\boldsymbol{\mu}$, while Figure 3(b) compares the observed number of available men and women as reported in the 1970 U.S. Census with the endogenously determined number of available men and women computed in the model, $\{\mu_{i,0}, \mu_{0,j}\}$. These are the underlying data used to estimate the marital gains. In order to rationalize the observed marriage distribution $\boldsymbol{\mu}$, the computed endogenous $\{\mu_{i,0}, \mu_{0,j}\}$ is significantly larger than the reported number of available men and women from the 1970 U.S. Census. Nonetheless, the computed $\{\mu_{i,0}, \mu_{0,j}\}$ maintains many of the qualitative features of the observed data.

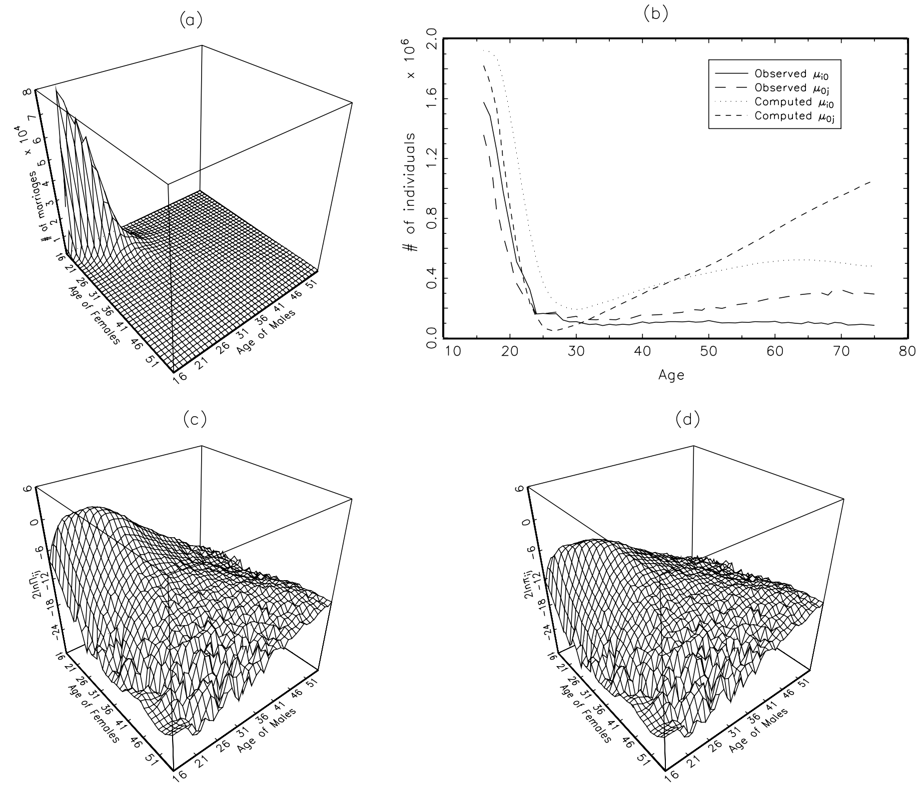

FIGURE 3.—(a) Observed new marriages in 1971/72. (b) Observed & computed μ_{i0}, μ_{0j} in 1971/72. (c) Dynamic Gains, $2\ln \Pi_{ij}$ for 1971/72. (d) Static Gains, $2\ln \pi_{ij}$ for 1971/72.

Figures 3(c) and 3(d) graph the estimates of the Dynamic Gains and Static Gains from marriage implied by this distribution of new marriages.[36] The Static Gains plot in Figure 3(d) shows strong assortative matching by age with the gains being highest for couples closest in age along the diagonal.[37] A peak occurs at an early age when young couples matched with each other. The plot of the Dynamic Gains in Figure 3(c) maintains many of the qualitative features of Figure 3(d). The Dynamic Gains plot shows a strong assortative matching pattern by age with the peak rising a lot higher and occurring at an even earlier age. Aside from the difference in the peak height, a portion of the Dynamic Gains plot for young age couples is now positive compared to Figure 3(d) where the Static Gains are all negative.

[36]The plots in Figures 3(c) and 3(d) use a nonparametric estimate to predict the gains for those age pairs where no marriages are observed. This typically happens for matches with large age differentials, that is, when a young individual is matched with a much older individual.

[37]CS provides a smoothed version of the Static Gains plot in Figure 3(c).

The Static Gains for an (i, j) pair are computed by taking the natural log of the number of current (i, j) matches divided by the geometric averages of those age i men and age j women that chose to remain single. This statistic ignores dynamic considerations in terms of forgone future opportunities of participating in the marriage market if individuals remain single. The Dynamic Gains statistic compensates for this shortcoming by internalizing the forgone future marriage market opportunities in the gains calculation. It approximates the future value of participating in the marriage market by using the probabilities of remaining single in the future given by $\prod_{k=0}^{T_{i,j}} (p_{i+k,0} q_{0,j+k})^{(1/2)\beta^k \prod_{l=1}^{k} \varrho_l \rho_l (1-\delta_{i,j,k})}$. Young individuals have the greatest opportunity to participate in the marriage market, albeit as increasingly older individuals as they age. Given that most marriages occur when individuals are young, the implied Dynamic Marriage Gains accounting for the forgone future marriage market opportunities are much larger than the Static Gains. The Static Gains statistic in effect assumes that there is only one opportunity to match and that, in the future, agents will remain single with certainty. In other words, future probabilities of remaining single, $p_{i+k,0}$ and $q_{i+k,0}$, equal 1.

Figure 4 graphs various cross-sections of the 1971/72 Static and Dynamic Gains against the age of individuals' spouses on the horizontal axis. Figures 4(a) and 4(b) plot the marriage gains for females aged 18, 25, and 34 years old, and Figures 4(c) and 4(d) plot the gains for males of the same ages. The graphs also plot the bootstrap 95% confidence interval computed using the procedure described in the Supplemental Material.[38] This set of graphs provides a more detailed picture of the differences between the Static Gains 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 is young. The Dynamic Gains for an 18 and a 25 year old far exceed their corresponding Static Gains and is positive when the spouse is young. The tight bootstrap confidence interval also suggests that the gains estimates are most precisely estimated for young individuals where most of the data lie.

Figures 5, 6, and 7 attempt to document the inter-temporal changes in marriage gains over the period from 1970 to 1990. In Figure 5, I construct a simple difference in the Static Gains and Dynamic Gains of marriage from 1970 to 1980. Figure 5 plots various cross-sections of this difference against the age of individuals' spouses on the horizontal axis. Figures 5(a) and 5(b) plot the differences for males and Figures 5(c) and 5(d) plot the differences for females. Generally, all these plots suggest a fall in marriage gains over this decade. Using the Dynamic Gains, Figures 5(a) and 5(c) suggest that the drop in marriage gains is even larger than initially suggested by the Static Gains calculation. The plots for females in Figures 5(c) and 5(d) suggest that calculations using Static

[38]For age pairs where no matches were observed, a nonparametric 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.

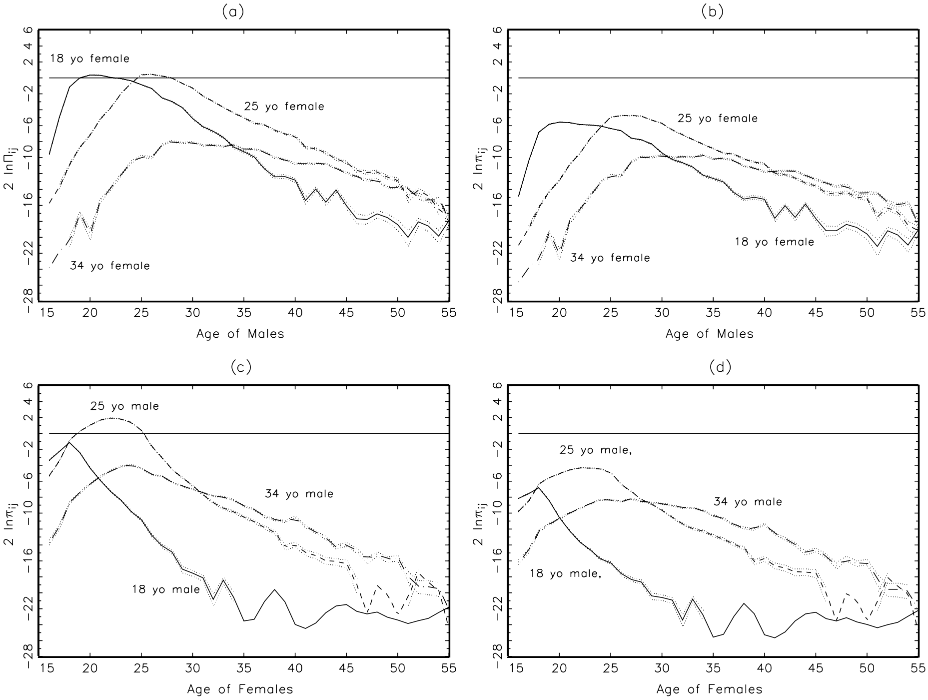

FIGURE 4.—(a) Dynamic Gains, $2\ln\Pi_{ij}$ for females. (b) Static Gains, $2\ln\pi_{ij}$ for females. (c) Dynamic Gains, $2\ln\Pi_{ij}$ for males. (d) Static Gains, $2\ln\pi_{ij}$ for males.

Gains understate the drop in marital gains across the entire range of the husband's age. The corresponding plots for males in Figures 5(a) and 5(b) suggest that the underestimate is mainly confined to the spouses' age younger than 30 years old.[39]

Figures 6(a), 6(b), and 6(c) present contour plots of the Dynamic Gains for 1971/72, 1981/82, and 1991/92, respectively. These plots trace out iso-level contour curves or locus of points that have equal level of Dynamic Gains. The curves in these plots are labeled with their corresponding levels of net utility. In general, the shape of these contour curves is in line with the qualitative features of the three-dimensional plot of the Dynamic Gains shown in Figure 3(b). The elliptical shape of the contours confirms the strong positive assortative matching pattern found in earlier plots. Couples that are closest in age realize the highest gains with the peak in Dynamic Gains occurring among couples that marry early.

Comparing the contour plots over the two decades, one also notices the changes in the range of contour levels over these periods. In Figure 6(a), the

[39]The plot of this marriage gains difference from 1980 to 1990 is qualitatively similar to Figure 5 and has been omitted.

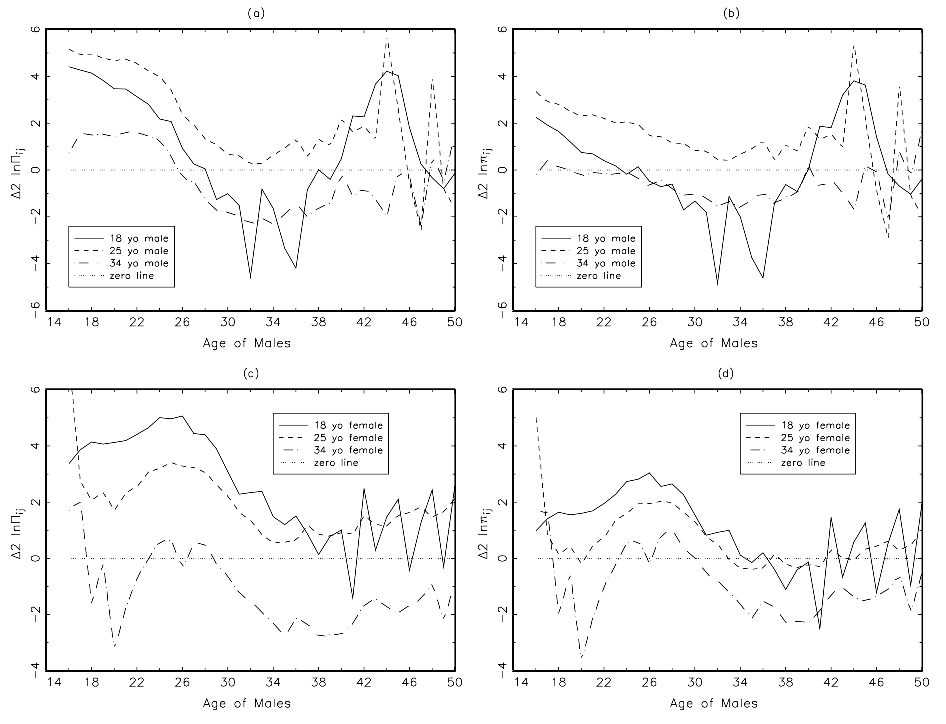

FIGURE 5.—(a) 70–80 dynamic diff.: $(2\ln \Pi_{ij}^{70} - 2\ln \Pi_{ij}^{80})$ for males. (b) 70–80 static diff.: $(2\ln \pi_{ij}^{70} - 2\ln \pi_{ij}^{80})$ for males. (c) 70–80 dynamic diff.: $(2\ln \Pi_{ij}^{70} - 2\ln \Pi_{ij}^{80})$ for females. (d) 70–80 static diff.: $(2\ln \pi_{ij}^{70} - 2\ln \pi_{ij}^{80})$ for females.

contour levels for the 1970 Dynamic Gains range in values from 0 to −15. This range of contour levels decreases to −3 to −15, in the contour plot for 1981/82 and 1991/92 in Figures 6(b) and 6(c), respectively. As the higher level contours disappear in the plots for 1981/82 and 1991/92, there is also a flattening of the marital gains with the peak moving away from the lower left corner toward the top right corner. All these are further confirmation of the drop in marriage gains among young couples highlighted in Figure 5, and the increase in the delay of marriage among young adults. Comparing the shape of the curves in Figures 6(a), 6(b), and 6(c), the curves have become more elliptical especially when comparing the 1971/72 and the 1981/82 contour plots. This is strongly suggestive of an increase in assortative matching by age over these periods.

Figure 7 investigates the marriage gains maximizing spousal age for men and women. These plots are meant to highlight the increase in assortative matching by age over the two decades from 1970 to 1990. Figure 7(a) plots the marriage gains maximizing spousal age for each female age over the two decades and Figure 7(c) repeats this for males. Figure 7(b) plots the difference between the marriage gains maximizing spousal age and the corresponding female age,

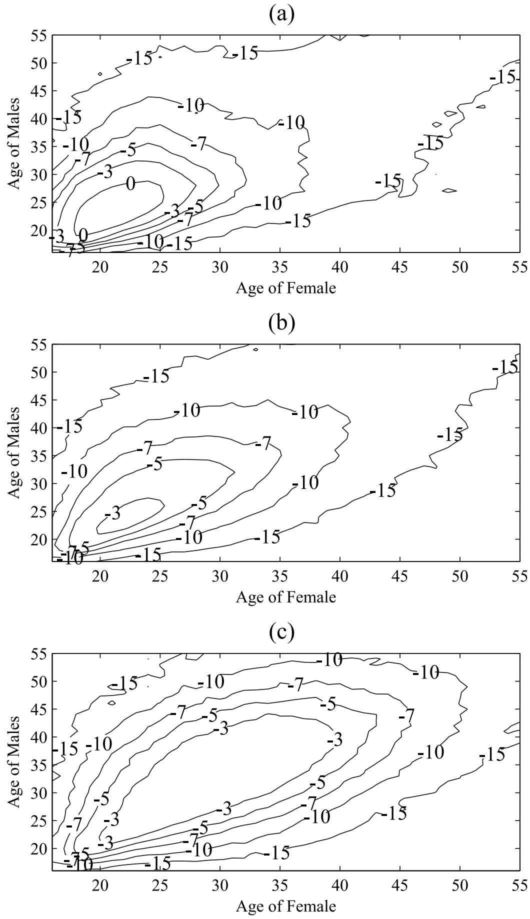

FIGURE 6.—(a) Contour of Dynamic Gains 71/72. (b) Contour of Dynamic Gains 81/82. (c) Contour of Dynamic Gains 91/92.

that is, it plots $(i_j^* - j)$ on the vertical axis and j on the horizontal axis, where i_j^* is the marriage gains maximizing spousal age for age j females. Similarly Figure 7(d) repeats this for males, that is, it plots $(j_i^* - i)$ against i. Focusing first on the plots for females, Figure 7(a) suggests that the marriage gains maximizing spousal age for females is very close to the corresponding female's age. Looking at Figure 7(b), women below the age of 25 years old maximize their Dynamic Gains when matched with slightly younger men, while women between the ages of 25 and 45 years old maximize their marriage gains when matched with slightly older men. Over the two decades, Figures 7(a) and 7(b) suggest that this age difference has decreased where the marriage gains maximizing spousal age is closer to the corresponding female's age. The plots for

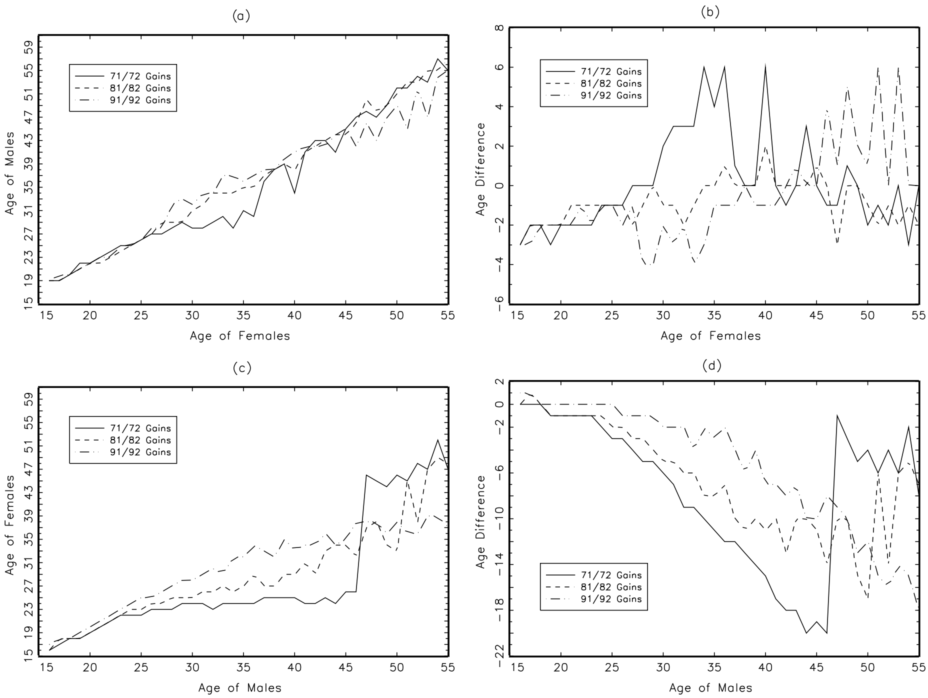

FIGURE 7.—(a) Dyn. Gains maximizing spousal age for females. (b) Dyn. Gains maximizing spousal age difference for females. (c) Dyn. Gains maximizing spousal age for males. (d) Dyn. Gains maximizing spousal age difference for males.

males in Figures 7(c) and 7(d) tell a slightly different story. They suggest that men's Dynamic Gains are maximized when matched with considerably younger women. For example, the estimates in 1971/72 suggest that 25 year old men (on average) maximize their marriage gains when matched with 21 year old women, 30 year old men with 21 year old women, 35 year old men with 24 year old women, and so on. This age difference appears to increase with older men. The 1971/72 estimates suggest a big decrease in age difference after approximately 47 years old, where 50 year old men (on average) maximize their marriage gains when matched with 46 year old women.[40] Comparing the 1971/72 estimates with those of 1981/82 and 1991/92, there is a clear decrease in the age difference between the marriage gains maximizing spousal age and the corresponding male's age. For example, by 1991/92 the estimates suggest that 25 year old men (on average) maximize their marriage gains when matched with 23 year old women, 30 year old men with 27 year old women, 35 year old men with 29 year old women, and so on.

[40]Given the sparseness of new marriages after the age of 45, these estimates should be interpreted with caution.

6. CONCLUSION

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

APPENDIX

A.1. *State Transition Matrices,* $\mathcal{F}_a(i' \mid i)$ *and* $\mathcal{R}_a(j' \mid j)$

$\mathcal{F}_a(i' \mid i)$ (or $\mathcal{R}_a(j' \mid j)$) denotes the transition probability that an age i male g (or an age j female h) will next find himself (or herself) single at age i' (or j') given his (or her) action a at age i (or j). If g chooses to be single, $a = 0$, then $\mathcal{F}_0(i+1 \mid i) = 1$ and $\mathcal{F}_0(r \mid i) = 0$ where $r \neq i+1$. Clearly, $r > i$ and all a.[41] That is, if g forgoes the opportunity to match at age i, he will be single with certainty in the next period at age $i+1$. Similarly, if an age $i < Z$ male matches with a female in her terminal age Z, he will return to the marriage market in the next period at age $i+1$ with certainty. That is, if $i < Z$ and $a = Z$, then $\mathcal{F}_Z(i+1 \mid i) = 1$. Consider now if g (of age $i \ll Z$) chooses to match with an older spouse of age j (where $i < j \ll Z$). The marriage might dissolve in the first period and g returns to the marriage market at age $i+1$. This occurs with probability $\mathcal{F}_j(i+1 \mid i) = \delta$. The marriage might survive the first period but dissolve in the second period, in which case he finds himself single again at age $i+2$. This occurs with probability $\mathcal{F}_j(i+2 \mid i) = \delta(1-\delta)$. The marriage could survive until the death of the older spouse and the younger g re-enters the marriage market at age $i + T_{i,j} + 1$. This occurs with probability $\mathcal{F}_j(i + T_{i,j} + 1 \mid i) = (1-\delta)^{T_{i,j}}$. The transition probability $\mathcal{F}_a(r \mid i)$ for $a \neq 0$ takes the form:
for $a = j \leq i$ (i.e., g marries someone younger),

$$\mathcal{F}_a(r \mid i) = \begin{cases} \delta(1-\delta)^{r-(i+1)}, & \text{if } i+1 \leq r \leq Z, \\ 0, & \text{elsewhere,} \end{cases}$$

for $a = j > i$ (g marries someone older),

$$\mathcal{F}_a(r \mid i) = \begin{cases} \delta(1-\delta)^{r-(i+1)}, & \text{if } i+1 \leq r \leq i + T_{i,j}, \\ (1-\delta)^{T_{i,j}}, & \text{if } r = i + T_{i,j} + 1, \\ 0, & \text{elsewhere.} \end{cases}$$

[41]That is, whatever the actions an age i male chooses, he can never return as a single age i or younger self in the future.

There is a similar analogous structure for the female transition matrix. $\mathcal{R}_a(r \mid j) = 0$ for all $r \leq j$ and all a. If h chooses to remain single at age j, she returns to the marriage market in the next period as an older female aged $j+1$ with certainty. That is, if $a = 0$, $\mathcal{R}_0(j+1 \mid j) = 1$ and zero elsewhere. The structure of $\mathcal{R}_a(r \mid j)$ for $a \neq 0$ is as follows:
for $a = i \leq j$ (h marries a younger man),

$$\mathcal{R}_a(r \mid j) = \begin{cases} \delta(1-\delta)^{r-(j+1)}, & \text{if } j+1 \leq r \leq Z, \\ 0, & \text{elsewhere,} \end{cases}$$

for $a = i > j$ (h marries someone older),

$$\mathcal{R}_a(r \mid i) = \begin{cases} \delta(1-\delta)^{r-(j+1)}, & \text{if } j+1 \leq r \leq j + T_{i,j}, \\ (1-\delta)^{T_{i,j}}, & \text{if } r = j + T_{i,j} + 1, \\ 0, & \text{elsewhere.} \end{cases}$$

A.2. *Derivation of Equations (3.18) and (3.19)*

McFadden (1978) introduced the family of choice models based on the Generalized Extreme Value Distribution, with distribution function

$$\text{(A.1)} \qquad \mathfrak{F}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_K) = \exp\bigl(-\mathcal{H}\bigl(e^{-\varepsilon_1}, e^{-\varepsilon_2}, \dots, e^{-\varepsilon_K}\bigr)\bigr).$$

$\mathcal{H}(\cdot)$ is a member of the class of functions from $\Re_+^K \to \Re_+$ with the following properties:

1. homogeneous of degree 1,
2. $\lim_{r_j \to \infty} \mathcal{H}(r_1, \dots, r_j, \dots, r_K) = \infty$,
3. the first partials of $\mathcal{H}$ are positive, and all the distinct cross-partials of order k (e.g., $\partial^k \mathcal{H} / \partial r_i \cdots \partial r_\ell$ for $i \neq \cdots \neq \ell$ are all distinct) are nonpositive if k is even, and nonnegative if k is odd.

For the Type 1 Extreme Value distribution, $\mathcal{H}(Y_{i,0}, Y_{i1}, \dots, Y_{iZ}) = \sum_{j=0}^{Z} Y_{i,j}$. Theorem 1 of McFadden (1978) showed that for utilities of the form $u_{ia} = \tilde{v}_{ia} + \varepsilon_{iag}$, where the vector $\boldsymbol{\varepsilon}_i = (\varepsilon_{i,0}, \varepsilon_{i1}, \dots, \varepsilon_{iZ})$ is distributed $\mathfrak{F}$, then the probability that j is selected satisfies $\mathcal{P}_{i,j} = e^{\tilde{v}_{i,j}} \mathcal{H}_j(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \cdot, e^{\tilde{v}_{iZ}}) / \mathcal{H}(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \cdot, e^{\tilde{v}_{iZ}})$. I reproduce the arguments of the proof here:

$$\text{(A.2)} \qquad \mathcal{P}_{i,j} = \int \mathbb{1}\Bigl\{ j = \arg\max_{a \in \mathcal{D}} (\tilde{v}_{ia} + \varepsilon_{iag}) \Bigr\} \mathfrak{f}(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon}$$
$$= \int_{\varepsilon_{i,j}=-\infty}^{\infty} \mathfrak{F}_j(\tilde{v}_{i,j} - \tilde{v}_{i,0} + \varepsilon_{i,j}, \tilde{v}_{i,j} - \tilde{v}_{i1} + \varepsilon_{i,j}, \dots,$$
$$\varepsilon_{i,j}, \dots, \tilde{v}_{i,j} - \tilde{v}_{iZ} + \varepsilon_{i,j})\, d\varepsilon_{i,j}$$

$$
\begin{aligned}
&= \int_{\varepsilon_{i,j}=-\infty}^{\infty} e^{-\varepsilon_{i,j}} \mathcal{H}_j\big(e^{\tilde{v}_{i,0}-\varepsilon_{i,j}-\tilde{v}_{i,j}}, e^{\tilde{v}_{i1}-\varepsilon_{i,j}-\tilde{v}_{i,j}}, \ldots, \\
&\quad e^{-\varepsilon_{i,j}}, \ldots, e^{\tilde{v}_{iZ}-\varepsilon_{i,j}-\tilde{v}_{i,j}}\big) \\
&\quad \times \exp\big[-\mathcal{H}\big(e^{\tilde{v}_{i,0}-\varepsilon_{i,j}-\tilde{v}_{i,j}}, e^{\tilde{v}_{i1}-\varepsilon_{i,j}-\tilde{v}_{i,j}}, \ldots, \\
&\quad e^{-\varepsilon_{i,j}}, \ldots, e^{\tilde{v}_{iZ}-\varepsilon_{i,j}-\tilde{v}_{i,j}}\big)\big]\, d\varepsilon_{i,j} \\
&= \int_{\varepsilon_{i,j}=-\infty}^{\infty} e^{-\varepsilon_{i,j}} \mathcal{H}_j\big(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \cdot, e^{\tilde{v}_{iZ}}\big) \\
&\quad \times \exp\big[-e^{-\varepsilon_{i,j}-\tilde{v}_{i,j}} \mathcal{H}\big(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \ldots, e^{\tilde{v}_{iZ}}\big)\big]\, d\varepsilon_{i,j} \\
&= \frac{e^{\tilde{v}_{i,j}} \mathcal{H}_j\big(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \cdot, e^{\tilde{v}_{iZ}}\big)}{\mathcal{H}\big(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \cdot, e^{\tilde{v}_{iZ}}\big)},
\end{aligned}
$$

where $\mathcal{H}_j(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \cdot, e^{\tilde{v}_{iZ}}) = \frac{\partial \mathcal{H}(\cdot)}{\partial e^{\tilde{v}_{i,j}}}$. The fourth equality uses the homogeneity of degree 1 of $\mathcal{H}$ and consequent homogeneity of degree zero of $\mathcal{H}_j$.

The following is a special case of the proof of Lemma 1 in Arcidiacono and Miller (2008). Consider the integrated value function

$$
\begin{aligned}
\mathbf{V}_i &= \mathbb{E} V_{\boldsymbol{\alpha}}(i, \boldsymbol{\varepsilon}_g) \\
&= \int V_{\boldsymbol{\alpha}}(i, \boldsymbol{\varepsilon}_g)\, d\mathfrak{F}(\boldsymbol{\varepsilon}_g) \\
&= \int \max_{a \in \mathcal{D}} (\tilde{v}_{ia} + \varepsilon_{iag}) \mathfrak{f}(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon} \\
&= \sum_{j=0}^{Z} \mathcal{P}_{i,j}(\tilde{v}_{i,j} + \varepsilon_{i,jg}) \\
&= \sum_{j=0}^{Z} \mathcal{P}_{i,j}\tilde{v}_{i,j} + \sum_{j=0}^{Z} \int \mathbb{1}\Big\{ j = \arg\max_{a \in \mathcal{D}} (\tilde{v}_{ia} + \varepsilon_{iag}) \Big\} \varepsilon_{i,jg} \mathfrak{f}(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon}.
\end{aligned}
$$

Subtracting $\tilde{v}_{i,0}$ on both sides, I get

$$
\begin{aligned}
\text{(A.3)} \qquad \mathbf{V}_i - \tilde{v}_{i,0} &= \sum_{j=0}^{Z} \mathcal{P}_{i,j}(\tilde{v}_{i,j} - \tilde{v}_{i,0}) \\
&\quad + \sum_{j=0}^{Z} \int \mathbb{1}\Big\{ j = \arg\max_{a \in \mathcal{D}} (\tilde{v}_{ia} + \varepsilon_{ia}) \Big\} \varepsilon_{i,j} \mathfrak{f}(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon}.
\end{aligned}
$$

Solving the second term,

$$\begin{aligned}
&\int \mathbb{1}\Big\{j = \arg\max_{a\in\mathcal{D}}(\tilde{v}_{ia} + \varepsilon_{iag})\Big\}\varepsilon_{i,jg}\mathfrak{f}(\boldsymbol{\varepsilon})\,d\boldsymbol{\varepsilon} \\
&\quad = \int_{\varepsilon_{i,j}=-\infty}^{\infty} \varepsilon_{i,j}\mathfrak{F}_j(\tilde{v}_{i,j} - \tilde{v}_{i,0} + \varepsilon_{i,j}, \tilde{v}_{i,j} - \tilde{v}_{i1} + \varepsilon_{i,j}, \dots, \\
&\qquad \varepsilon_{i,j}, \dots, \tilde{v}_{i,j} - \tilde{v}_{iZ} + \varepsilon_{i,j})\,d\varepsilon_{i,j} \\
&\quad = \mathcal{H}_j\big(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \dots, e^{\tilde{v}_{iZ}}\big)\int_{\varepsilon_{i,j}=-\infty}^{\infty} \varepsilon_{i,j}e^{-\varepsilon_{i,j}} \\
&\qquad \times \exp\big[-e^{-\varepsilon_{i,j}-\tilde{v}_{i,j}}\mathcal{H}\big(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \dots, e^{\tilde{v}_{iZ}}\big)\big]\,d\varepsilon_{i,j} \\
&\quad = \mathcal{H}_j\big(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \dots, e^{\tilde{v}_{iZ}}\big)\frac{e^{e^{\tilde{v}_{i,j}}}\big(\ln\mathcal{H}(\cdot) - \tilde{v}_{i,j} + c\big)}{\mathcal{H}\big(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \dots, e^{\tilde{v}_{iZ}}\big)} \\
&\quad = \mathcal{P}_{i,j}\big(\ln\mathcal{H}(\cdot) - \tilde{v}_{i,j} + c\big).
\end{aligned}$$

Using the property of the mean of a Type 1 Extreme Value distribution where

$$\int_{x=-\infty}^{\infty} \frac{x}{\phi}\exp\big(-e^{-(x-\xi)/\phi}\big)e^{-(x-\xi)/\phi}\,dx = \xi + c\phi,$$

where $c \approx 0.57722$,

$$\text{(A.4)}\qquad \begin{aligned}
&\mathbf{V}_i - \tilde{v}_{i,0} = \sum_{j=0}^{Z}\mathcal{P}_{i,j}(\tilde{v}_{i,j} - \tilde{v}_{i,0}) + \sum_{j=0}^{Z}\mathcal{P}_{i,j}\big(\ln\mathcal{H}(\cdot) - \tilde{v}_{i,j} + c\big) \\
&\quad\Rightarrow\quad \mathbf{V}_i = \ln\mathcal{H}(\cdot) + c.
\end{aligned}$$

In the case where $\boldsymbol{\varepsilon}_i$ is i.i.d., Type 1 Extreme Value,

$$\mathcal{P}_{i,0} = \frac{e^{\tilde{v}_{i,0}}}{\mathcal{H}\big(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \dots, e^{\tilde{v}_{iZ}}\big)}.$$

Taking logs on both sides,

$$\ln\mathcal{H}\big(e^{\tilde{v}_{i,0}}, e^{\tilde{v}_{i1}}, \dots, e^{\tilde{v}_{iZ}}\big) = \tilde{v}_{i,0} - \ln\mathcal{P}_{i,0},$$

and substituting this into Equation (A.4), I get

$$\mathbf{V}_i = \tilde{v}_{i,0} + c - \ln\mathcal{P}_{i,0}.$$

Consider a single individual in his terminal age Z, from Equation (3.13) $\tilde{v}_{i,0} = \alpha_{i,0}$. This gives me Equation (3.18) when $i = Z$,

$$\mathbf{V}_i = \alpha_{i,0} + c - \ln\mathcal{P}_{i,0}.$$

Consider next when $i < Z$, accordingly the mean utility from being single as given by Equation (3.13):

$$\tilde{v}_{i,0} = \alpha_{i,0} + \sum_{k=i+1}^{Z} \beta^{k-i}\mathbb{E}\big[V_{\boldsymbol{\alpha}}(k, \boldsymbol{\varepsilon}_{kg}) \mid i, \boldsymbol{\varepsilon}_{ig}, a_{ig} = 0\big].$$

The CI assumption allows us to factorize the expectation $\mathbb{E}[V_{\boldsymbol{\alpha}}(k, \boldsymbol{\varepsilon}_{k,g}) \mid i, \boldsymbol{\varepsilon}_{i,g}, a_{i,g} = 0] = \mathcal{F}_0(k \mid i)\int[V_{\boldsymbol{\alpha}}(k, \boldsymbol{\varepsilon}_{k,g})\, dF(\boldsymbol{\varepsilon}_{k,g})$. Since $\mathcal{F}_0(i+1 \mid i) = 1$ and 0 elsewhere, $\sum_{k=i+1}^{Z} \beta^{k-i}\mathbb{E}[V_{\boldsymbol{\alpha}}(k, \boldsymbol{\varepsilon}_{k,g}) \mid i, \boldsymbol{\varepsilon}_{i,g}, a_{i,g} = 0] = \sum_{k=i+1}^{Z} \beta^{k-i}\mathcal{F}_0(k \mid i)\int[V_{\boldsymbol{\alpha}}(k, \boldsymbol{\varepsilon}_{k,g})\, d\mathfrak{F}(\boldsymbol{\varepsilon}_{k,g}) = \beta\mathbf{V}_{i+1}$. Hence $\tilde{v}_{i,0} = \alpha_{i,0} + \beta\mathbf{V}_{i+1}$. Substituting this into Equation (A.4), I get Equation (3.18) when $i < Z$,

$$\mathbf{V}_i = \alpha_{i,0} + c + \beta\mathbf{V}_{i+1} - \ln\mathcal{P}_{i,0}.$$

A.3. *Derivation of Equation (3.20)*

The proof proceeds by backward induction. It comprises two steps. I first establish that Equation (3.20) holds for $i = Z$. In the second step, I show that when Equation (3.20) holds for any age i, then it also holds for age $i - 1$.

Step 1: If $i = Z$, the match duration for any (Z, j) marriage is $T_{Z,j} = Z - \max(Z, j) + 1 = 1$. As defined in Equation (3.13), $\tilde{v}_{Z,j} = \boldsymbol{\alpha}_Z(j) - \tau_{Z,j}$. Hence Equation (3.20) holds for $i = Z$.

Step 2: I consider two cases.

Case 1: When i marries a younger spouse, that is, $j < i-1$ and $T_{i,j} = Z-i+1$. I want to show that if

$$\tilde{v}_{i,j} = \boldsymbol{\alpha}_i(j) - \tau_{i,j} + \sum_{k=1}^{T_{i,j}-1} \beta^k\big(1 - (1-\delta)^k\big)(\alpha_{i+k,0} + c - \ln\mathcal{P}_{i+k,0}),$$

then,

$$\begin{aligned}\tilde{v}_{i-1,j} &= \boldsymbol{\alpha}_{i-1}(j) - \tau_{i-1,j}\\ &\quad + \sum_{k=1}^{T_{i-1,j}-1} \beta^k\big(1 - (1-\delta)^k\big)(\alpha_{i-1+k,0} + c - \ln\mathcal{P}_{i-1+k,0}).\end{aligned}$$

Since $j < i-1$, then $T_{i-1,j} = Z - (i-1) + 1 = T_{i,j} + 1$. The choice specific value function, $\tilde{v}_{i,j}$ as defined in Equation (3.13) can be expressed as:

$$\begin{aligned}\tilde{v}_{i,j} = \boldsymbol{\alpha}_i(j) - \tau_{i,j} &+ \beta\mathbb{E}\big[V_{\boldsymbol{\alpha}}(i+1, \boldsymbol{\varepsilon}_{i+1,g}) \mid i, \boldsymbol{\varepsilon}_{i,g}, a_{i,g} = j\big]\\ &+ \beta^2\mathbb{E}\big[V_{\boldsymbol{\alpha}}(i+2, \boldsymbol{\varepsilon}_{i+2,g}) \mid i, \boldsymbol{\varepsilon}_{i,g}, a_{i,g} = j\big] + \cdots\\ &+ \beta^{T_{i,j}-2}\mathbb{E}\big[V_{\boldsymbol{\alpha}}(i + T_{i,j} - 2, \boldsymbol{\varepsilon}_{i+T_{i,j}-2,g}) \mid i, \boldsymbol{\varepsilon}_{i,g}, a_{i,g} = j\big]\end{aligned}$$

$$
\begin{aligned}
&+ \beta^{T_{i,j}-1}\mathbb{E}\big[V_{\boldsymbol{\alpha}}(i+T_{i,j}-1, \boldsymbol{\varepsilon}_{i+T_{i,j}-1,g}) \mid i, \boldsymbol{\varepsilon}_{i,g}, a_{i,g}=j\big] \\
= {}& \boldsymbol{\alpha}_i(j) - \tau_{i,j} + \beta\mathcal{F}_j(i+1 \mid i)\int V_{\boldsymbol{\alpha}}(i+1, \boldsymbol{\varepsilon}_{i+1,g})\, d\mathfrak{F}(\boldsymbol{\varepsilon}_{i+1,g}) \\
&+ \beta^2\mathcal{F}_j(i+2 \mid i)\int V_{\boldsymbol{\alpha}}(i+2, \boldsymbol{\varepsilon}_{i+2,g})\, d\mathfrak{F}(\boldsymbol{\varepsilon}_{i+2,g}) + \cdots \\
&+ \beta^{T_{i,j}-2}\mathcal{F}_j(i+T_{i,j}-2 \mid i) \\
&\times \int V_{\boldsymbol{\alpha}}(i+T_{i,j}-2, \boldsymbol{\varepsilon}_{i+T_{i,j}-2,g})\, d\mathfrak{F}(\boldsymbol{\varepsilon}_{i+T_{i,j}-2,g}) \\
&+ \beta^{T_{i,j}-1}\mathcal{F}_j(i+T_{i,j}-1 \mid i) \\
&\times \int V_{\boldsymbol{\alpha}}(i+T_{i,j}-1, \boldsymbol{\varepsilon}_{i+T_{i,j}-1,g})\, d\mathfrak{F}(\boldsymbol{\varepsilon}_{i+T_{i,j}-1,g}).
\end{aligned}
$$

By the definition of the state transition probabilities given in Appendix A.1, and the integrated value function for a single age i male $\mathbf{V}_i = \mathbb{E}V_{\boldsymbol{\alpha}}(i, \boldsymbol{\varepsilon}_g) = \int V_{\boldsymbol{\alpha}}(i, \boldsymbol{\varepsilon}_g)\mathfrak{f}(d\boldsymbol{\varepsilon}_g)$,

$$
\begin{aligned}
\tilde{v}_{i,j} = {}& \boldsymbol{\alpha}_i(j) - \tau_{i,j} + \beta\delta\mathbf{V}_{i+1} + \beta^2\delta(1-\delta)\mathbf{V}_{i+2} + \cdots \\
&+ \beta^{T_{i,j}-2}\delta(1-\delta)^{T_{i,j}-3}\mathbf{V}_{i+T_{i,j}-2} + \beta^{T_{i,j}-1}\delta(1-\delta)^{T_{i,j}-2}\mathbf{V}_{i+T_{i,j}-1}.
\end{aligned}
$$

By repeated substitution of Equation (3.18) starting from $\mathbf{V}_{i+1}$ to $\mathbf{V}_{i+T_{i,j}-1}$, we get

$$
\begin{aligned}
\tilde{v}_{i,j} = {}& \boldsymbol{\alpha}_i(j) - \tau_{i,j} + \beta\delta(\alpha_{i+1,0} + c - \ln\mathcal{P}_{i+1,0}) \\
&+ \beta^2\delta\big(1+(1-\delta)\big)(\alpha_{i+2,0} + c - \ln\mathcal{P}_{i+2,0}) + \cdots \\
&+ \beta^k\delta\sum_{s=0}^{k-1}(1-\delta)^s(\alpha_{i+k,0} + c - \ln\mathcal{P}_{i+k,0}) + \cdots \\
&+ \beta^{T_{i,j}-2}\delta\sum_{s=0}^{T_{i,j}-3}(1-\delta)^s(\alpha_{i+T_{i,j}-2,0} + c - \ln\mathcal{P}_{i+T_{i,j}-2,0}) \\
&+ \beta^{T_{i,j}-1}\delta\sum_{s=0}^{T_{i,j}-2}(1-\delta)^s(\alpha_{i+T_{i,j}-1,0} + c - \ln\mathcal{P}_{i+T_{i,j}-1,0}).
\end{aligned}
$$

Since $(1-\delta) < 1$, the geometric series $\sum_{s=0}^{k-1}(1-\delta)^s = (1-(1-\delta)^k)/\delta$. Hence we arrive at

$$
\tilde{v}_{i,j} = \big(\boldsymbol{\alpha}_i(j) - \tau_{i,j}\big) + \sum_{k=1}^{T_{i,j}-1}\beta^k\big(1-(1-\delta)^k\big)(\alpha_{i+k,0} + c - \ln\mathcal{P}_{i+k,0}). \tag{A.5}
$$

Using the definition in Equation (3.13) for $\tilde{v}_{i-1,j}$ as earlier,

$$\begin{aligned}\tilde{v}_{i-1,j} = \boldsymbol{\alpha}_{i-1}(j) - \tau_{i-1,j} + \beta\delta\mathbf{V}_i + \beta^2\delta(1-\delta)\mathbf{V}_{i+1} + \cdots \\ + \beta^{T_{i-1,j}-2}\delta(1-\delta)^{T_{i-1,j}-3}\mathbf{V}_{i-1+T_{i-1,j}-2} \\ + \beta^{T_{i-1,j}-1}\delta(1-\delta)^{T_{i-1,j}-2}\mathbf{V}_{i-1+T_{i-1,j}-1}.\end{aligned}$$

Substituting in Equation (3.18),

$$\begin{aligned}\tilde{v}_{i-1,j} &= \boldsymbol{\alpha}_{i-1}(j) - \tau_{i-1,j} + \beta\delta(\alpha_{i,0} + c - \ln\mathcal{P}_{i,0}) + \beta(\beta\delta\mathbf{V}_{i+1}) \\ &+ \beta^2\delta(1-\delta)(\alpha_{i+1,0} + c - \ln\mathcal{P}_{i+1,0}) + \beta\big(\beta^2\delta(1-\delta)\mathbf{V}_{i+2}\big) \\ &+ \cdots \\ &+ \beta^{T_{i-1,j}-2}\delta(1-\delta)^{T_{i-1,j}-3}(\alpha_{i-1+T_{i-1,j}-2,0} + c - \ln\mathcal{P}_{i-1+T_{i-1,j}-2,0}) \\ &+ \beta\big(\beta^{T_{i-1,j}-2}\delta(1-\delta)^{T_{i-1,j}-3}\mathbf{V}_{i-1+T_{i-1,j}-1}\big) \\ &+ \beta^{T_{i-1,j}-1}\delta(1-\delta)^{T_{i-1,j}-2}(\alpha_{i-1+T_{i-1,j}-1,0} + c - \ln\mathcal{P}_{i-1+T_{i-1,j}-1,0}).\end{aligned} \tag{A.6}$$

Since $T_{i,j} = Z - i + 1 = T_{i-1,j} - 1$, then $T_{i-1,j} - 2 = T_{i,j} - 1$ and so on. From Equation (A.5),

$$\begin{aligned}\tilde{v}_{i,j} - \big(\boldsymbol{\alpha}_i(j) - \tau_{i,j}\big) &= \beta\delta\mathbf{V}_{i+1} + \beta^2\delta(1-\delta)\mathbf{V}_{i+2} + \cdots \\ &+ \beta^{T_{i,j}-2}\delta(1-\delta)^{T_{i,j}-3}\mathbf{V}_{i+T_{i,j}-2} \\ &+ \beta^{T_{i,j}-1}\delta(1-\delta)^{T_{i,j}-2}\mathbf{V}_{i+T_{i,j}-1}.\end{aligned} \tag{A.7}$$

Substituting Equation (A.7) into Equation (A.6),

$$\begin{aligned}\tilde{v}_{i-1,j} &= \big(\boldsymbol{\alpha}_{i-1}(j) - \tau_{i-1,j}\big) + \beta\big(\tilde{v}_{i,j} - \big(\boldsymbol{\alpha}_i(j) - \tau_{i,j}\big)\big) \\ &+ \beta\delta(\alpha_{i,0} + c - \ln\mathcal{P}_{i,0}) \\ &+ \beta^2\delta(1-\delta)(\alpha_{i+1,0} + c - \ln\mathcal{P}_{i+1,0}) + \cdots \\ &+ \beta^{T_{i-1,j}-2}\delta(1-\delta)^{T_{i-1,j}-3}(\alpha_{i-1+T_{i-1,j}-2,0} + c - \ln\mathcal{P}_{i-1+T_{i-1,j}-2,0}) \\ &+ \beta^{T_{i-1,j}-1}\delta(1-\delta)^{T_{i-1,j}-2}(\alpha_{i-1+T_{i-1,j}-1,0} + c - \ln\mathcal{P}_{i-1+T_{i-1,j}-1,0}).\end{aligned}$$

Plugging in Equation (A.5), I get

$$\begin{aligned}\tilde{v}_{i-1,j} &= \big(\boldsymbol{\alpha}_{i-1}(j) - \tau_{i-1,j}\big) + \beta\delta(\alpha_{i,0} + c - \ln\mathcal{P}_{i,0}) \\ &+ \beta^2\big[\big(1-(1-\delta)\big) + \delta(1-\delta)\big](\alpha_{i+1,0} + c - \ln\mathcal{P}_{i+1,0}) \\ &+ \beta^3\big[\big(1-(1-\delta)^2\big) + \delta(1-\delta)^2\big](\alpha_{i+2,0} + c - \ln\mathcal{P}_{i+2,0}) \\ &+ \cdots\end{aligned}$$

$$
\begin{aligned}
& + \beta^{T_{i-1,j}-2}\big[\delta(1-\delta)^{T_{i-1,j}-3} + \delta(1-\delta)^{T_{i-1,j}-3}\big] \\
& \times (\alpha_{i-1+T_{i-1,j}-2,0} + c - \ln \mathcal{P}_{i-1+T_{i-1,j}-2,0}) \\
& + \beta^{T_{i-1,j}-1}\big[\delta(1-\delta)^{T_{i-1,j}-2} + \delta(1-\delta)^{T_{i-1,j}-2}\big] \\
& \times (\alpha_{i-1+T_{i-1,j}-1,0} + c - \ln \mathcal{P}_{i-1+T_{i-1,j}-1,0}) \\
\Rightarrow \quad & \tilde{v}_{i-1,j} = \big(\boldsymbol{\alpha}_{i-1}(j) - \tau_{i-1,j}\big) \\
& \qquad + \sum_{k=1}^{T_{i-1,j}-1} \beta^k\big(1-(1-\delta)^k\big)(\alpha_{i-1+k,0} + c - \ln \mathcal{P}_{i-1+k,0}).
\end{aligned}
$$

Case 2: When i marries a spouse at least as old as himself, that is, $j \geq i$ and $T_{i,j} = Z - j + 1$. I want to show that if

$$
\begin{aligned}
\tilde{v}_{i,j} = {} & \boldsymbol{\alpha}_i(j) - \tau_{i,j} \\
& + \sum_{k=1}^{T_{i,j}-1} \beta^k\big(1-(1-\delta)^k\big)(\alpha_{i+k,0} + c - \ln \mathcal{P}_{i+k,0}) \\
& + \sum_{k=i+T_{i,j}}^{Z} \beta^{k-i}(\alpha_{k,0} + c - \ln \mathcal{P}_{k,0}),
\end{aligned}
$$

then,

$$
\begin{aligned}
\tilde{v}_{i-1,j} = {} & \boldsymbol{\alpha}_{i-1}(j) - \tau_{i-1,j} \\
& + \sum_{k=1}^{T_{i-1,j}-1} \beta^k\big(1-(1-\delta)^k\big)(\alpha_{i-1+k,0} + c - \ln \mathcal{P}_{i-1+k,0}) \\
& + \sum_{k=i-1+T_{i-1,j}}^{Z} \beta^{k-(i-1)}(\alpha_{k,0} + c - \ln \mathcal{P}_{k,0}).
\end{aligned}
$$

Using the definition in Equation (3.13):

$$
\begin{aligned}
\tilde{v}_{i,j} = {} & \boldsymbol{\alpha}_i(j) - \tau_{i,j} + \beta\delta\mathbf{V}_{i+1} + \beta^2\delta(1-\delta)\mathbf{V}_{i+2} + \cdots \\
& + \beta^{T_{i,j}-1}\delta(1-\delta)^{T_{i,j}-2}\mathbf{V}_{i+T_{i,j}-1} + \beta^{T_{i,j}}(1-\delta)^{T_{i,j}-1}\mathbf{V}_{i+T_{i,j}}.
\end{aligned} \tag{A.8}
$$

By repeated substitution of Equation (3.18), I get

$$
\begin{aligned}
\tilde{v}_{i,j} = {} & \boldsymbol{\alpha}_i(j) - \tau_{i,j} + \beta\delta(\alpha_{i+1,0} + c - \ln \mathcal{P}_{i+1,0}) \\
& + \beta^2\big(1-(1-\delta)^2\big)(\alpha_{i+2,0} + c - \ln \mathcal{P}_{i+2,0}) + \cdots
\end{aligned} \tag{A.9}
$$

$$
\begin{aligned}
&+\beta^{T_{i,j}-2}\big(1-(1-\delta)^{T_{i,j}-2}\big)(\alpha_{i+T_{i,j}-2,0}+c-\ln\mathcal{P}_{i+T_{i,j}-2,0})\\
&+\beta^{T_{i,j}-1}\big(1-(1-\delta)^{T_{i,j}-1}\big)(\alpha_{i+T_{i,j}-1,0}+c-\ln\mathcal{P}_{i+T_{i,j}-1,0})\\
&+\beta^{T_{i,j}}\big(1-(1-\delta)^{T_{i,j}-1}+(1-\delta)^{T_{i,j}-1}\big)\\
&\times(\alpha_{i+T_{i,j},0}+c-\ln\mathcal{P}_{i+T_{i,j},0})+\cdots\\
&+\beta^{Z-i}\big(1-(1-\delta)^{T_{i,j}-1}+(1-\delta)^{T_{i,j}-1}\big)(\alpha_{Z,0}+c-\ln\mathcal{P}_{Z,0})
\end{aligned}
$$

$$
\text{(A.10)}\quad \Rightarrow\quad \tilde{v}_{i,j}=\boldsymbol{\alpha}_i(j)-\tau_{i,j}+\sum_{k=1}^{T_{i,j}-1}\beta^k\big(1-(1-\delta)^k\big)(\alpha_{i+k,0}+c-\ln\mathcal{P}_{i+k,0})+\sum_{k=i+T_{i,j}}^{Z}\beta^{k-i}(\alpha_{k,0}+c-\ln\mathcal{P}_{k,0}).
$$

Using the definition in Equation (3.13) for $\tilde{v}_{i-1,j}$ as earlier,

$$
\begin{aligned}
\tilde{v}_{i-1,j}=\boldsymbol{\alpha}_{i-1}(j)-\tau_{i-1,j}&+\beta\delta\mathbf{V}_i+\beta^2\delta(1-\delta)\mathbf{V}_{i+1}+\cdots\\
&+\beta^{T_{i-1,j}-1}\delta(1-\delta)^{T_{i-1,j}-2}\mathbf{V}_{i-1+T_{i-1,j}-1}\\
&+\beta^{T_{i-1,j}}(1-\delta)^{T_{i-1,j}-1}\mathbf{V}_{i-1+T_{i-1,j}}.
\end{aligned}
$$

Substituting in Equation (3.18),

$$
\text{(A.11)}\quad
\begin{aligned}
\tilde{v}_{i-1,j}=\boldsymbol{\alpha}_{i-1}(j)-\tau_{i-1,j}&+\beta\delta(\alpha_{i,0}+c-\ln\mathcal{P}_{i,0})+\beta(\beta\delta\mathbf{V}_{i+1})\\
&+\beta^2\delta(1-\delta)(\alpha_{i+1,0}+c-\ln\mathcal{P}_{i+1,0})+\beta\big(\beta^2\delta(1-\delta)\mathbf{V}_{i+2}\big)\\
&+\cdots\\
&+\beta^{T_{i-1,j}-1}\delta(1-\delta)^{T_{i-1,j}-2}(\alpha_{i-1+T_{i-1,j}-1,0}+c-\ln\mathcal{P}_{i-1+T_{i-1,j}-1,0})\\
&+\beta\big(\beta^{T_{i-1,j}-1}\delta(1-\delta)^{T_{i-1,j}-2}\mathbf{V}_{i-1+T_{i-1,j}}\big)\\
&+\beta^{T_{i-1,j}}(1-\delta)^{T_{i-1,j}-1}(\alpha_{i-1+T_{i-1,j},0}+c-\ln\mathcal{P}_{i-1+T_{i-1,j},0})\\
&+\beta\big(\beta^{T_{i-1,j}}(1-\delta)^{T_{i-1,j}-1}\mathbf{V}_{i-1+T_{i-1,j}+1}\big).
\end{aligned}
$$

Since $j\geq i$, then $T_{i,j}-1=Z-j=T_{i-1,j}-1$, and similarly $T_{i-1,j}=T_{i,j}=Z-j+1$. Substituting in Equation (A.8),

$$
\text{(A.12)}\quad
\begin{aligned}
\tilde{v}_{i-1,j}=\boldsymbol{\alpha}_{i-1}(j)-\tau_{i-1,j}&+\beta\delta(\alpha_{i,0}+c-\ln\mathcal{P}_{i,0})\\
&+\beta\big(\tilde{v}_{i,j}-\big(\boldsymbol{\alpha}_i(j)-\tau_{i,j}\big)\big)\\
&+\beta^2\delta(1-\delta)(\alpha_{i+1,0}+c-\ln\mathcal{P}_{i+1,0})\\
&+\cdots
\end{aligned}
$$

$$+ \beta^{T_{i-1,j}-1}\delta(1-\delta)^{T_{i-1,j}-2}(\alpha_{i-1+T_{i-1,j}-1,0} + c - \ln \mathcal{P}_{i-1+T_{i-1,j}-1,0})$$
$$+ \beta^{T_{i-1,j}}(1-\delta)^{T_{i-1,j}-1}(\alpha_{i-1+T_{i-1,j},0} + c - \ln \mathcal{P}_{i-1+T_{i-1,j},0}).$$

Substituting in (A.11),

$$\begin{aligned}
\tilde{v}_{i-1,j} &= \big(\boldsymbol{\alpha}_{i-1}(j) - \tau_{i-1,j}\big) + \beta\big(1-(1-\delta)\big)(\alpha_{i,0} + c - \ln \mathcal{P}_{i,0}) \\
&\quad + \beta^2\big[\big(1-(1-\delta)\big) + \delta(1-\delta)\big](\alpha_{i-1+2,0} + c - \ln \mathcal{P}_{i-1+2,0}) \\
&\quad + \beta^3\big[\big(1-(1-\delta)^2\big) + \delta(1-\delta)^2\big](\alpha_{i-1+3,0} + c - \ln \mathcal{P}_{i-1+3,0}) \\
&\quad + \cdots \\
&\quad + \beta^{T_{i-1,j}-1}\big[\big(1-(1-\delta)^{T_{i-1,j}-2}\big) + \delta(1-\delta)^{T_{i-1,j}-2}\big] \\
&\quad \times (\alpha_{i-1+T_{i-1,j}-1,0} + c - \ln \mathcal{P}_{i-1+T_{i-1,j}-1,0}) \\
&\quad + \beta^{T_{i-1,j}}\big[\big(1-(1-\delta)^{T_{i-1,j}-1}\big) + (1-\delta)^{T_{i-1,j}-1}\big] \\
&\quad \times (\alpha_{i-1+T_{i-1,j},0} + c - \ln \mathcal{P}_{i-1+T_{i-1,j},0}) \\
&\quad + \sum_{k=i-1+T_{i-1,j}+1}^{Z} \beta^{k-(i-1)}(\alpha_{k,0} + c - \ln \mathcal{P}_{k,0}) \\
\Rightarrow \quad \tilde{v}_{i-1,j} &= \big(\boldsymbol{\alpha}_{i-1}(j) - \tau_{i-1,j}\big) \\
&\quad + \sum_{k=1}^{T_{i-1,j}-1} \beta^k\big(1-(1-\delta)^k\big)(\alpha_{i-1+k,0} + c - \ln \mathcal{P}_{i-1+k,0}) \\
&\quad + \sum_{k=i-1+T_{i-1,j}}^{Z} \beta^{k-(i-1)}(\alpha_{k,0} + c - \ln \mathcal{P}_{k,0}).
\end{aligned}$$

This completes the derivation of Equation (3.20).

The derivation of Equation (3.21) is straightforward: Starting from the definition of $\tilde{v}_{i,0}$ in Equation (3.13), I repeatedly substitute for all $\mathbf{V}_k$ according to Equation (3.18). After some simple algebra, I arrive at (3.21). The derivations of the female side Equations (3.23) and (3.23) are mirror images of the derivation presented above and have been omitted.

A.4. *Derivation of Equation (3.24)*

From Equations (3.20), (3.21), and (3.16),

$$\begin{aligned}
\ln\left(\frac{\mathcal{P}_{i,j}}{\mathcal{P}_{i,0}}\right) &= \tilde{v}_{i,j} - \tilde{v}_{i,0} \\
&= \boldsymbol{\alpha}_i(j) - \tau_{i,j} - \alpha_{i,0}
\end{aligned}$$

$$- \sum_{k=1}^{T_{i,j}-1} \big(\beta(1-\delta)\big)^k \big(\alpha_{i+k,0} + c + \ln \mathcal{P}_{i+k,0}^{-1}\big)$$

$$= \boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_{i,0}(j) - \tau_{i,j} - \sum_{k=1}^{T_{i,j}-1} \big(\beta(1-\delta)\big)^k \big(c + \ln \mathcal{P}_{i+k,0}^{-1}\big),$$

where

$$\boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_{i,0}(j) = \sum_{k=0}^{T_{i,j}-1} \big(\beta(1-\delta)\big)^{k-1} (\alpha_{i,j,k} - \alpha_{i+k,0}).$$

The derivation of Equation (3.26) mirrors the step just outlined.

A.5. *Estimation of the Availables, m_1 and f_1*

The model assumes that the flow of new marriages comes from a stationary data generating process that is in steady state. In each period, m_1 and f_1 men and women enter the system; their lifespans are determined by the gender-specific mortality rates ($\boldsymbol{\rho}, \boldsymbol{\varrho}$). The flow of new marriages is driven by the marital gains $\boldsymbol{\Pi}$ and the number of available men and women $\mathbf{m}$ and $\mathbf{f}$. The available men and women at ages $(i > 1, j > 1)$ come from

1. agents who chose to remain single at ages $(i-1, j-1)$ and survived to age (i, j),
2. the flow of previously married age $(i-1, j-1)$ agents whose marriage dissolved either due to death of a spouse or divorce governed by $\boldsymbol{\delta}$.

The model also assumes that the marriage market is closed with new entrants only entering at age 1 which will not hold when confronted with data.

Setting m_1 and f_1 to the levels reported in the U.S. Census for 1970, 1980, and 1980, however, fails the dynamic constraints in Equations (5.4) and (5.5) generating levels of available men and women that are not able to sustain the flow of new marriages. In allowing supplies to be endogenous, one is faced with the challenge of estimating the level of m_1 and f_1 entering the system that satisfy the dynamic constraints while maintaining the relative differences of available men and women observed in the data over the three decades. I take the approach of jointly estimating the smallest triple, $(\hat{m}_{1,70}, \hat{m}_{1,80}, \hat{m}_{1,90})$, $(\hat{f}_{1,70}, \hat{f}_{1,80}, \hat{f}_{1,90})$, such that the estimated $\hat{m}_{1,t}$ and $\hat{f}_{1,t}$ maintain the relative difference observed in data while satisfying the dynamic constraints for their respective time period.[42] The estimates are given in Table A-I. The observed supplies of male and females are graphed in Figures A1(a) and A1(b), while the estimated supplies are graphed in Figures A1(c) and A1(d).

[42]That is, $\hat{m}_{1,70} - \hat{m}_{1,80} = m_{1,70} - m_{1,80}$, $\hat{m}_{1,70} - \hat{m}_{1,90} = m_{1,70} - m_{1,90}$, $\hat{m}_{1,70} - \hat{f}_{1,70} = m_{1,70} - f_{1,70}$, $\hat{f}_{1,70} - \hat{f}_{1,80} = f_{1,70} - f_{1,80}$, and $\hat{f}_{1,70} - \hat{f}_{1,90} = f_{1,70} - f_{1,90}$, where $m_{1,t}$ and $\hat{m}_{1,t}$ are observed and estimated male quantities, respectively, for year t.

TABLE A-I

ESTIMATES OF $\hat{m}_1$ AND $\hat{f}_1$

Years	1970	1980	1990
$\hat{m}_1$ ($\times 10^6$)	1.922	2.025	1.790
$\hat{f}_1$ ($\times 10^6$)	1.819	1.941	1.719

The estimated supply of available men and women for 1970 and 1980 seems to maintain much of the qualitative features observed in the data. The estimated supplies for these two years also capture some of the differences we see in data. While the model under-predicts the difference between 1970 and 1980 for the young ages, the estimated difference seems largest for those aged between 25 and 30 years old as observed in the data. The model also systematically over-predicts the number of available men and women at the later ages. The model seems to have the most difficulty matching the supplies in 1990. The estimated supplies for 1990 fail to replicate much of the qualitative features we see in the data.

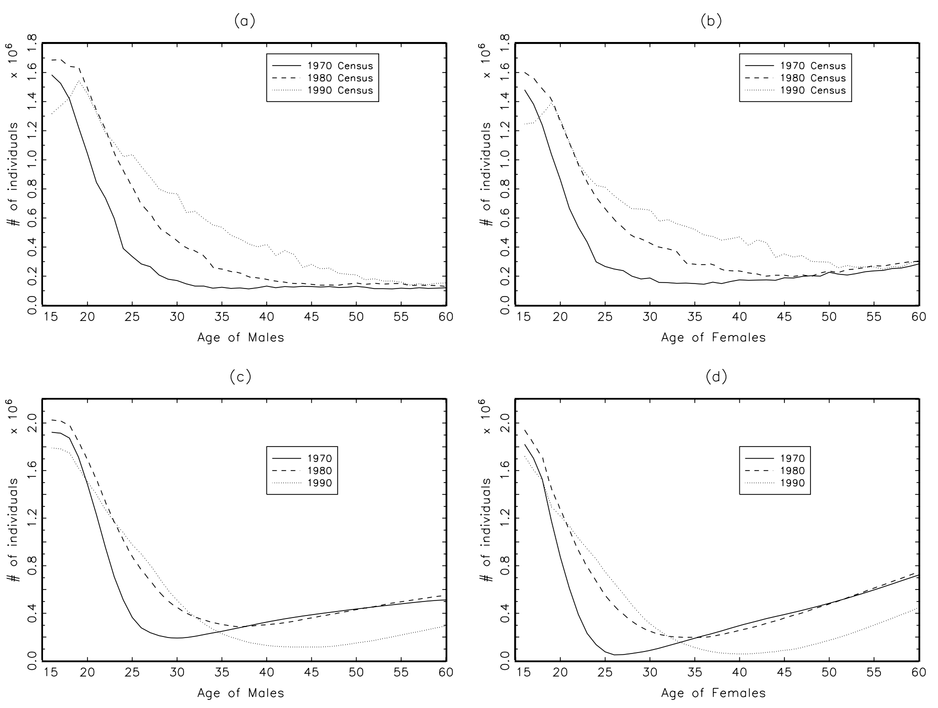

FIGURE A1.—(a) Single males from Census data. (b) Single females from Census data. (c) Estimated single males. (d) Estimated single females.

REFERENCES

AGRAWAL, N. (2013): "An Empirical Model of the Medical Match," Report, MIT. [1376]

AMUEDO-DORANTES, C., AND S. GROSSBARD-SCHECTMAN (2007): "Cohort Level Sex Ratio Effects on Women's Labor Force Participation," *Review of Economics of the Household*, 5 (3), 249–278. [1377]

ARCIDIACONO, P., AND R. A. MILLER (2008): "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," Working Paper, Tepper School. [1412]

AZEVEDO, E. M., AND J. W. HATFIELD (2013): "Complementarity and Multidimensional Heterogeneity in Large Matching Markets," Report, University of Pennsylvania. [1392]

BERRY, S. (1994): "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242–262. [1376]

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63 (4), 841–890. [1376]

BRESNAHAN, T. (1987): "Competition and Collusion in the American Automobile Oligopoly: The 1955 Price War," *Journal of Industrial Economics*, 35 (4), 457–482. [1376]

BRIEN, M. J., L. A. LILLARD, AND S. STERN (2006): "Cohabitation, Marriage, and Divorce in a Model of Match Quality," *International Economic Review*, 47 (2), 451–494. [1376,1377]

BRUZE, G., M. SVARER, AND Y. WEISS (2015): "The Dynamics of Marriage and Divorce," *Journal of Labor Economics*, 33 (1), 123–170. [1377]

CHEN, L. (2013): "Managerial Talent Misallocation and the Cost of Moral Hazard," Report, University of Calgary. [1376]

CHIAPPORI, P.-A. (1988): "Rational Household Labor Supply," *Econometrica*, 56 (1), 63–90. [1377]

——— (1992): "Collective Labor Supply and Welfare," *Journal of Political Economy*, 100 (3), 437–467. [1377]

CHIAPPORI, P.-A., B. FORTIN, AND G. LACROIX (2002): "Household Labor Supply, Sharing Rule, and the Marriage Market," *Journal of Political Economy*, 110 (1), 37–72. [1377]

CHIAPPORI, P.-A., B. SALANIÉ, AND Y. WEISS (2011): "Partner Choice and the Marital College Premium," Report, Columbia University. [1376]

CHOO, E. (2015): "Supplement to 'Dynamic Marriage Matching: An Empirical Framework'," *Econometrica Supplemental Material*, 83, http://dx.doi.org/10.3982/ECTA10675. [1374,1375]

CHOO, E., AND S. SEITZ (2013): "The Collective Matching Model: Identification, Estimation and Testing," in *Structural Econometric Models*. Advances in Econometrics, Vol. 31, ed. by E. Choo and M. Shum. Cambridge: Emerald Publishing Group Limited. [1377]

CHOO, E., AND A. SIOW (2006): "Who Marries Whom and Why," *Journal of Political Economy* 114 (1), 175–201. [1374]

——— (2007): "Lifecycle Marriage Matching: Theory and Evidence," Working Paper, University of Toronto. [1377,1378]

DEBREU, G. (1960): "Review of R. D. Luce Individual Choice Behavior," *American Economic Review*, 50, 186–188. [1376]

DECKER, C., E. H. LIEB, R. J. MCCANN, AND B. K. STEPHENS (2013): "Unique Equilibria and Substitution Effects in a Stochastic Model of the Marriage Market," *Journal of Economic Theory*, 148, 778–792. [1376]

ECHENIQUE, F., S. LEE, M. SHUM, AND M. B. YENMEZ (2013): "The Revealed Preference Theory of Stable and Extremal Stable Matchings," *Econometrica*, 81, 153–171. [1376]

FOX, J. T. (2010): "Estimating Matching Games With Transfers," Working Paper, University of Michigan. [1376]

GALICHON, A., AND B. SALANIÉ (2012): "Cupid's Invisible Hands: Social Surplus and Identification in Matching Models," Report, Columbia University. [1376]

GRAHAM, B. S. (2011): "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers," in *Handbook of Social Economics*, Vol. 1B, ed. by J. Benhabib, A. Bisin, and M. O. Jackson. Amsterdam: North-Holland. [1376]

——— (2013): "Comparative Static and Computational Methods for an Empirical One-to-One Transferable Utility Matching Model," in *Structural Econometric Models*. Advances in Econometrics, Vol. 31, ed. by E. Choo and M. Shum. Cambridge: Emerald Publishing Group Limited, 151–179. [1376]

HITSCH, G., A. HORTAÇSU, AND D. ARIELY (2010): "Matching and Sorting in Online Dating Markets," *American Economic Review*, 100 (1), 130–163. [1377]

HSIEH, Y.-W. (2013): "Understanding Mate Preference From Two-Sided Matching Markets: Identification, Estimation and Policy Analysis," Report, University of Southern California. [1377]

IYIGUN, M. (2009): "Marriage, Cohabitation and Commitment," Report, University of Colorado. [1397]

LEE, S., AND M. NIEDERLE (2011): "Propose With a Rose? Signaling in Internet Dating Markets," Report, University of Maryland. [1377]

LUNDBERG, S., AND R. POLLAK (1996): "Bargaining and Distribution in Marriage," *Journal of Economic Perspectives*, 10 (4), 139–158. [1377]

MAZZOCCO, M., C. RUIZ, AND S. YAMAGUCHI (2007): "Labor Supply, Wealth Dynamics, and Marriage Decisions," Report, University of California, Los Angeles. [1397]

MCFADDEN, D. (1974): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, ed. by P. Zarembka. New York: Academic Press. [1376]

——— (1978): "Modelling the Choice of Residential Location," in *Spatial Interaction and Planning Models*, ed. by F. Snickars and J. Weibull. Amsterdam: North-Holland. [1376,1411]

MENZEL, K. (2015): "Large Matching Markets as Two Sided Demand Systems," *Econometrica*, 83 (3), 897–941. [1377]

PETRIN, A. (2002): "Quantifying the Benefits of New Products: The Case of Minivan," *Journal of Political Economy*, 110 (4), 705–729. [1376]

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55 (5), 999–1033. [1374,1377,1383,1387]

——— (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, Vol. 4, ed. by R. E. Engle and D. McFadden. Amsterdam: Elsevier–North-Holland, Chapter 14, 3081–3143. [1384]

SAUTMANN, A. (2011): "Age Matching Patterns and Search," Report, Brown University. [1380]

SEITZ, S. (2009): "Accounting for Racial Differences in Marriage and Employment," *Journal of Labor Economics*, 27 (3), 385–437. [1377]

SHIMER, R., AND L. SMITH (2000): "Assortative Matching and Search," *Econometrica*, 68 (2), 343–369. [1380]

VAN DER KLAAUW, W. (1996): "Female Labour Supply and Marital Status Decisions: A Life-Cycle Model," *Review of Economic Studies*, 63 (2), 199–235. [1377]

WEISS, Y. (1997): "The Formation and Dissolution of Families: Why Marry? Who Marries Whom? And What Happens Upon Divorce," in *Handbook of Population and Family Economics*, ed. by M. Rosenzweig and O. Stark. Amsterdam: North-Holland. [1383]

WEISS, Y., AND R. WILLIS (1997): "Match Quality, New Information and Marital Dissolution," *Journal of Labor Economics*, 15, s293–s329. [1397]

Dept. of Economics, University of Calgary, 2500 University Drive NW, Calgary, AB, T2N1N4, Canada; echoo@ucalgary.ca.

Manuscript received March, 2012; final revision received December, 2014.