

Cohabitation, Marriage, and Divorce in a Model of Match Quality*

Michael J. Brien Lee A. Lillard
Deloitte & Touche LLP University of Michigan

Steven Stern
University of Virginia

September 2002

Abstract

The objective of this research is to develop and estimate an economic model of nonmarital cohabitation, marriage, and divorce that is consistent with current data on the formation and dissolution of relationships. Jovanovic's (1979) theoretical matching model is extended to help explain household formation and dissolution behavior. Implications of the model reveal what factors influence the decision to start a relationship, what form this relationship will take, and the relative stability of the various types of unions. The structural parameters of the model are estimated using longitudinal data from a sample of female high school seniors from the U.S. New

*We would like to thank B. Ravikumar, William Johnson, Francis Kramarz, Bob Rosenthal, and seminar participants at the 1997 Summer Research Workshop at the Institute for Research on Poverty, University of Michigan, University of Virginia, Virginia Tech, George Washington University, the University of North Carolina at Greensboro, Boston University, the 1998 Economic Demography Workshop at the meetings of Population Association of America, the 1998 SITE meetings, and the Southern Economic Association meetings. This research was partially supported by Grants P-50-12639, R01-HD30856, and R01-HD31585 from the NICHD. Brien also would like to acknowledge research support from the Center for Children, Families, and the Law at the University of Virginia. The views expressed in this paper are those of the authors and should not be attributed to the institutions with which they are affiliated. All mistakes are our own.

numerical methods are developed to reduce computational costs associated with estimation. The empirical results are mostly consistent with previous literature but have interesting interpretations given the structural model.

1. Introduction

When Gary Becker wrote his first work on the theory of marriage (Becker 1973), the world was a much different place. Since then, the United States has seen increases in the age at first marriage and a dramatic rise in the rate of divorce. It has been estimated, for example, that over half of all first marriages will be disrupted (Cherlin 1992). One of the most glaring changes since Becker's early work, however, has been the increase in the number of couples living together outside of a formal marriage. Bumpass and Sweet (1989) report that, for women born between 1940 and 1944, only 3% had ever lived in a nonmarital cohabitation by age 25. For women born 20 years later, 37% had cohabited by that same age. The evidence actually indicates that, despite the increases in the age of first marriage, individuals are still forming coresidential relationships at about the same point in their lives (Bumpass, Sweet, and Cherlin 1991). Even though they are frequently converted into marriages, nonmarital cohabiting relationships are typically much different from traditional marriages. They are shorter lived and, in general, cohabiting couples exhibit different behavior than married couples in such things as the investment in relationship specific capital (e.g. children).¹ Yet, by the same token, cohabiting relationships are different from being single. They represent a form of commitment between partners that may be a stepping stone to a marriage.

In light of these large demographic changes, the overall objective of this research is to develop and estimate an economic model of nonmarital cohabitation, marriage, and divorce that is consistent with current data on the formation and dissolution of relationships. Since Becker's pioneering work, research on the economics of the family has focused more on how individuals deal with their uncertain futures, on the dynamics of behavior, and on the interrelationship between decisions.² For example, Becker, Landes, and Michael (1977) examine the issue of

¹See Rindfuss and VandenHeuval (1990) for a comparison of cohabiting and marriage relationships.

²Becker (1973) acknowledges the existence of nonmarital cohabitation and attempts to make his theory sufficiently general to include this form of relationship but does not distinguish between the two forms. In fact, for his theoretical work, "marriage" is defined to mean that two

marital instability and the role of uncertainty. Building on this work, Weiss and Willis (1997) examine the role of match quality and the evolution of information within marital relationships. Becker (1973, 1991) has suggested an interaction between cohabitation and marriage decisions by arguing that the potential for marital instability can lead to “trial marriages.” The model proposed here directly follows an extensive literature in labor economics on search and decision making in the presence of uncertainty (e.g. Jovanovic 1979, Harris and Weiss 1984, Miller 1984). An important feature is the recognition that gathering information about a potential partner is a key aspect of the courtship process. Information learned during courtship influences not only the likelihood but also the manner in which relationships are formed. Acquiring information about a partner does not end with the formation of a union; it continues throughout the relationship where new information may influence the subsequent stability of that relationship. This research attempts to formalize these ideas by explicitly modeling the information gathering process within a relationship. This research is consistent with the notion of “intensive search” discussed by Becker, Landes, and Michael (1977) and the later work by Mortensen (1988) that applies search theory to many of these same decisions.

The model presented in this paper is the same type as developed by Jovanovic (1979) in which match quality is an “experience” good. In that paper, he focuses on the interactions between a worker and a firm when a worker’s productivity is not immediately known to the firm. One significant difference between that model and the current one is that individuals can now choose an intermediate type of relationship, a nonmarital cohabitation. Our model reveals what factors influence the decision to start a relationship, what form this relationship will take, and the relative stability of various types of unions. This allows us to address, for example, how changes in the cost of divorce due to reforms in the legal environment influence the stability of marriages and the manner in which relationships are formed.

After developing the theoretical framework and exploring some of its implications, we use data from the National Longitudinal Study of the High School Class of 1972 (NLS72) to estimate the structural parameters of the model. Estimation is accomplished by a modification of maximum simulated likelihood (MSL) and is in the spirit of an ordered discrete choice model. An innovation developed for this paper involves interpolation and weighting methods to significantly reduce

individuals “share the same household.” See Weiss (1997) for a recent survey of the literature on household formation and dissolution.

computation costs. The parameter estimates are used both to gauge the validity of the theoretical model and to simulate a variety of policy and “comparative dynamics” experiments. The structural nature of the estimation procedure allows for straightforward interpretation of estimation and simulation results.

The remainder of this paper proceeds as follows. In the next section, we describe the stylized facts that capture the current demographic environment. Section 3 describes the theoretical model. Section 4 derives some of the properties of the model, focussing on which results in this class of models are sensitive to strong or arbitrary assumptions. Section 5 describes the data. Section 6 presents the estimation methods. Section 7 presents the structural estimates and accompanying simulations. Section 8 concludes.

2. Stylized Facts

Prior to developing the model, we first describe the transition rates in and out of relationships the model should replicate. To do this, data from the nationally representative NLS72 is used. This data, which is described in more detail below, followed a relatively large group of men and women from the time they were high school seniors in 1972 until 1986. The 1986 interview included a retrospective history about the timing of cohabitation, marriage, marital disruption, and childbearing. This history includes the starting and ending dates of up to three relationships; cohabitations and marriages. The NLS72 is well suited for the issues addressed in this research because the large sample is representative of all young adults enrolled in their senior year of high school and because of the rich, longitudinal information available. It is important to note that most of, if not all, the following observations have been made in other papers using the same or comparable data.³

1. Many relationships begin as a nonmarital cohabitation

In slightly under 30% of the cases in the NLS72, the first time a white respondent lived with someone was outside of a formal marriage. A slightly larger fraction of black respondents began their first relationship in this manner. Entry into cohabitation becomes more prevalent when the respondent is older.

³Willis and Michael (1994) use the NLS72, Bumpass and Sweet (1989) use the National Survey of Families and Households, and Lauman et al. (1994) use the National Health and Social Life Survey.

2. Cohabitations are shorter-lived than marriages

The duration of coresidential relationships by the type of the union is examined in Panel A of Figure 1. This figure shows Kaplan-Meier survivor estimates for first marriages and first cohabitations by the race of the respondent. For whites, 97% of first marriages remain by the first anniversary. In contrast, only about 50% of the cohabitations survived the first year. By the third anniversary, approximately 80% of marriages and 20% of cohabitations were intact. While a similar pattern is revealed for blacks, their first marriages are less stable and their first cohabitations are more stable.

3. Many cohabitations are converted into marriages

There are three possible ways that our data on nonmarital cohabitations can end: the relationship ends, the cohabitation is converted into a marriage, or the observation is censored (i.e., the cohabitation was still active at the end of the NLS72 panel). Seventy-three percent of white cohabitations and 60% of black cohabitations end with the partners marrying. Overall, approximately 20% of first nonmarital cohabitations end with the partners separating.

4. The risk of separation declines with the duration of the relationship

Panel B of Figure 1 shows the hazard of separation from cohabitations and marriages for white and black respondents. After a rise over the first two years, the hazard of separation declines over the remainder of the panel. Consistent with the survivor functions presented in Panel A, black marriages appear less stable. A similar declining risk of separation in labor market data on job turnover provided much of the motivation for Jovanovic's (1979) earlier study.

5. The risk of divorce in marriages preceded by a nonmarital cohabitation with the same partner is higher than in other marriages

Evidence from the NLS72 as well as other data sources suggest that couples who choose to live together prior to the start of a formal marriage face an increased risk of marital dissolution (Bennett, Blanc and Bloom 1988; DeMaris and Rao 1992; Teachman and Polonko 1990). This result is somewhat counter-intuitive if one believes that only the cohabiting relationships with the greatest chance for success are converted into a formal marriage. Lillard, Brien and Waite

(1995) present evidence of unobserved heterogeneity in the decisions to cohabit and divorce, and that the basic finding that marriages preceded by cohabitation having greater instability is due to self-selection into cohabitation.

In addition to these basic facts, a large body of literature has examined the many other factors that influence when and if a person chooses to start or end a coresidential relationship. The empirical results presented below examine the role of some of these factors - race, religion, education, and children - and the remainder of this section briefly summarizes some of the key facts in this literature. A number of papers have considered the influence of race on family structure decisions (see, for example, Brien 1997). As shown extensively in the literature and demonstrated in the figures above, blacks exhibit vastly different behaviors than whites with regard to living arrangements. This may be due in part to the resources available from the partners of black women or to the opportunities available to black women while single. Other research has considered the role of religion and found, for example, that Catholics marry later than non-Catholics (e.g., Michael and Tuma 1985). This effect may be due to Catholic restrictions on divorce.⁴ Researchers have shown that a higher level of education and the associated longer period of school enrollment lead to a delay in family formation (e.g., Brien, Lillard and Waite 1999). A possible explanation for this may be greater employment opportunities for more highly educated women outside of a relationship. It might also indicate that students do not have a life-style or schedule that is conducive to marriage or that they risk losing financial support from their parents if they enter a coresidential relationship while enrolled in school (e.g., Thornton, Axinn, and Teachman 1995). Finally, the literature on family structure has considered the impact of children on marriage and cohabitation decisions. The evidence in this area suggests, on the one hand, that individuals form relationships immediately after becoming pregnant or having a child, suggesting that a child precipitates the start of a coresidential relationship. This effect, however, appears to be fleeting in that the presence of children born outside of a coresidential relationship can eventually lead to a lower hazard of marriage (e.g., Bennett, Bloom, and Miller 1995; Upchurch, Lillard, and Panis 1999). On the other hand, the evidence also suggests that individuals have increased risk of having a child after entering into a coresidential relationship. Overall, the linkage between marriage and childbearing

⁴One might conjecture that the prohibition on contraceptive use reduces sexual activity and, therefore, the benefits of being in certain types of relationships for Catholics relative to non-Catholics. But Laumann et al. (1994) find that Catholics are as likely to use contraception as non-Catholics and have sex as frequently as non-Catholics.

has weakened over time (Akerlof, Yellen, and Katz 1995).

3. Theoretical Model

3.1. General Motivation

In this section we construct a simple model of union formation and dissolution in the presence of uncertain match quality. The model presented in this paper is a discrete-time extension of the model originally formulated by Jovanovic (1979).⁵ Some of the modelling assumptions, however, are driven by a desire to estimate the model. Individuals are assumed to have a finite life and to maximize the discounted present value of utility. Single individuals (i.e., individuals not currently in a coresidential relationship) randomly meet one potential partner at the start of each period. If coupled with a particular partner, the resulting relationship would provide a level of match quality to each of the participants which is assumed not to change over the course of a relationship.⁶ The exact match quality associated with a particular partner is assumed not to be immediately observed. Rather, an individual receives a noisy signal of the true quality. Upon receiving a signal, an individual decides whether to start a relationship with that person or to continue searching for a new partner (i.e., wait until the following period and see what a new, randomly-drawn partner has to offer). If she chooses to start a relationship, she also must decide what form that relationship should take: non-marital cohabitation or marriage. The value of the relationship depends upon the signal rather than the unobserved match quality in that she derives direct utility from the signal and that it provides information about the distribution of future signals. The decision to form a relationship is made prior to the realization of the true match quality and, therefore, must be based upon the expectation of the true match quality conditional upon the initial draw.⁷

⁵See also the presentation in Sargent (1987).

⁶An alternative way of modeling would be to assume that the match quality stochastically changes over the duration of the relationship. Weiss and Willis (1985, 1997) and Drewianka (1998) model stochastic shocks to relationships in a way that is consistent with both learning and changes in match quality over time. Neither paper, however, estimates their model.

⁷In reality, of course, the two partners together typically make the decision to start and end a relationship (Becker, 1991). However, if we follow an earlier literature (Pissarides 1983; Dagsvik, Jovanovic, and Shepard 1985) and assume linear rent sharing rules, then it is safe to act as if the partner that we are not modeling is passive. A linear rent sharing rule results in a decentralized (non-cooperative) Nash bargain. Alternatively, we could just rely on the efficiency

The model allows there to be differences in the value associated with a non-marital cohabitation and a formal marriage conditional on match quality. First, it is assumed that the utility flow from a given match is different for a cohabitation than a marriage. It may be the case, for instance, that individuals receive an additional utility bonus to being married relative to being in a cohabitation. When combined with other assumptions, if this bonus did not exist, everyone would prefer to cohabit. This utility bonus can be thought of, for example, as a gain to the division of labor within the household that may be particular to marriage or a lower price (or greater legal protections) associated with raising children within marriage (see Weiss 1997). Second, it is assumed that there is a cost associated with dissolving each type of relationship, and this cost may be different for a marriage versus a cohabitation. This cost can be thought of as the legal costs of obtaining a divorce or the psychic costs associated with a break-up.

The uncertainty about the quality of the match is slowly stripped away after entering a relationship. The rate at which the partners learn about the match is assumed to be the same whether the relationship is a marriage or a cohabitation. After being in the relationship for one period and learning more about quality of the match, the individual can choose to change the status of the relationship. If married, the individual can choose to stay married or to divorce. If cohabiting, the individual can choose to continue cohabiting, convert the relationship into a formal marriage, or dissolve the relationship altogether. Each additional period that the person is in a relationship, she learns more about the quality of the match and makes a comparable set of decisions. If the agent ever chooses to become single, she remains single for the period and then receives a new match the next period.⁸ The process continues over the agent's life.

property used in Browning et al. (1994). This would imply that, whenever there are positive rents, they are split in a way so that both partners want to remain in the relationship. Then we would just have to assume that the wife's share of the rents had the distributional properties we assume.

⁸This assumes an intervening period of being single between coresidential relationships. In fact, in the raw NLS72 cohabitation and marital histories, when moving between a first and second coresidential relationship, only 3.5% of women reported the first relationship ending and the second relationship starting in the same month and year.

3.2. Basic Framework

3.2.1. Match Quality and Relationships

The relationship status of the agent at age t is denoted as $m_t = 1$ iff single, $m_t = 2$ iff cohabitating, and $m_t = 3$ iff married. The duration with any partner at age t is denoted by d_t . Let θ denote the unknown match quality. It is assumed that $\theta \sim N(0, \sigma_\theta^2)$.⁹ Each period, couples draw a noisy signal of this match quality, ε_t , which is observed by the agent and affects utility. It is assumed that, conditional on θ ,

$$\begin{aligned} \varepsilon_t - \theta &= \eta_t && \text{if } d_t = 1, \\ \varepsilon_t - \theta &= \rho(\varepsilon_{t-1} - \theta) + \eta_t && \text{if } d_t > 1, \\ \eta_t &\sim N\left(0, \frac{\sigma_\eta^2}{1 - \rho^2}\right) && \text{for } d_t = 1 \\ \eta_t &\sim iidN(0, \sigma_\eta^2) && \text{for } d_t > 1. \end{aligned} \tag{3.1}$$

The noisy signal, ε_t , is modeled as an AR(1) process with autocorrelation term $\rho > 0$ to allow for divorces late in a relationship that can not be explained by learning. That is, if ρ is large enough, then a bad draw of ε_t implies bad times for many periods to come even if θ is known with precision and is satisfactory. The serial correlation in ε_t also implies a correlation in the choices made by the agent while learning is occurring.¹⁰ We model η_t at the beginning of a relationship ($d_t = 1$) as having a larger variance than subsequent η_t 's so that the ε_t process is time-stationary with covariance matrix

$$\Omega = \frac{\sigma_\eta^2}{1 - \rho^2} \begin{pmatrix} 1 & \rho & \dots & \rho^{d_t-1} \\ \rho & 1 & \dots & \rho^{d_t-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{d_t-1} & \rho^{d_t-2} & \dots & 1 \end{pmatrix}.$$

Throughout, we assume that the agent knows ρ , σ_η^2 , and σ_θ^2 .

Each period that the agent is in a relationship, she receives a new signal of the quality of the match. Using the new signal, she updates the prediction of the true

⁹There is no loss of generality in assuming that $E\theta = 0$ because $E\theta$ will not be identified from the constant in a flow equation introduced below.

¹⁰The theoretical results in this paper can be generalized by allowing the vector of realizations $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots)$ to be affiliated as defined in Milgrom and Weber (1982). We impose normality throughout, however, to simplify presentation and to implement estimation.

match quality. Equation (3.1) implies a Bayesian estimate, $\hat{\theta}_{d_t}$, for θ at t using the last d_t observations of the noisy signal, $\underline{\varepsilon}'_{d_t} = (\varepsilon_{t+1-d_t} \ \varepsilon_{t+2-d_t} \ \dots \ \varepsilon_t)$.¹¹ It can be shown that $\hat{\theta}_{d_t}$ is a weighted average of $E\theta (= 0)$ and the generalized least squares (GLS) estimator of $\hat{\theta}_{d_t}$ where the weights are respectively σ_θ^{-2} and the reciprocal of the variance of the GLS estimator.¹² This allows us to treat $\hat{\theta}_{d_t}$ and ε_t as state variables.

3.2.2. Children and Other Covariates

The model contains a number of other components that are believed to influence these decisions. Let X_t be the set of exogenous variables including a constant, age, race, and education. Each covariate in X_t is limited to exogenous characteristics with known paths.¹³

An element of family structure clearly important in the formation and dissolution of relationships is children. Given the relative complexity of the model, it is very difficult (computationally expensive) to incorporate childbearing decisions as one of the primary choices made by the agents. However, rather than treat childbearing decisions as simply an exogenous covariate, we allow the probability that the agent conceives a child to be a function of personal characteristics. Let $c_t =$ child-related covariates [$c_{1t} =$ number of children, $c_{2t} =$ age of youngest child]. Let $b_t = 1$ iff a child is conceived in t .¹⁴ The probability of a child being conceived in period t is assumed logistic, so that

$$p_t = e^{\pi_t} / [1 + e^{\pi_t}] \quad (3.2)$$

where π_t varies with relationships; that is¹⁵

$$\pi_t = \gamma_0(m_t) + c_t \gamma_c(m_t) + d_t \gamma_d(m_t) + t \gamma_t(m_t) + X_t \gamma_x(m_t). \quad (3.3)$$

3.2.3. Upper Bounds on State Variables

The state space for the dynamic programming problem implied by the model above is too large without imposing some upper bounds on some of the state variables. In particular, assume that no more learning occurs after some relationship

¹¹Note that a relationship of length d_t at age t started at age $t + 1 - d_t$.

¹²See Appendix A in Brien, Lillard, and Stern (2000) for details.

¹³ X_t includes a piecewise linear spline in age.

¹⁴No child deaths are allowed.

¹⁵In the empirical implementation, we allow age t , duration d_t , and child characteristics c_t to affect π_t nonlinearly through spline functions.

duration \bar{t}_d . Thus $\hat{\theta}_{d_t} = \hat{\theta}_{d_{t-1}}$ for all $d_t \geq \bar{t}_d$. One can think of what happens at \bar{t}_d as either a revelation similar to the learning process in Sargent (1987),¹⁶ or the updating equation for $\hat{\theta}_{d_t}$, equation (3) in Brien, Lillard and Stern (2000), is well approximated by ignoring learning after \bar{t}_d .¹⁷ Also, duration effect is allowed after \bar{t}_d . Thus, we censor d_t at \bar{t}_d , and the updating rule for d_t is

$$d_{t+1} = \min(d_t + 1, \bar{t}_d) 1(m_{t+1} = m_t \cup (m_{t+1} = 3 \cap m_t = 2)) + 1(m_{t+1} \neq m_t \cap (m_{t+1} \neq 3 \cup m_t \neq 2)) \quad (3.4)$$

where $1(\bullet)$ is an indicator function equal to one iff its argument is true.

Similarly, it is necessary to limit the range of the child characteristics. Let \bar{n}_c be the maximum number of children and \bar{n}_{ca} be the maximum age of the youngest child relevant to the value functions. The implication of these limits is not that people can not have more than \bar{n}_c children or a youngest child older than \bar{n}_{ca} ; rather it just implies that extra children beyond \bar{n}_c and the age of a child greater than \bar{n}_{ca} have no marginal effect on the value function. Thus, the updating rules for c_t are

$$\begin{aligned} c_{1t+1} &= \min(c_{1t} + b_t, \bar{n}_c), \\ c_{2t+1} &= (1 - b_t) \min(c_{2t} + 1, \bar{n}_{ca}). \end{aligned} \quad (3.5)$$

The value of these assumptions is that the related state variables are truncated at these upper bounds.

3.3. Value Function

We now characterize the utility associated with each of the potential choices available to the agent. Let $S_t = (m_t, m_{t-1}, c_{t-1}, d_t, \hat{\theta}_{d_t}, \varepsilon_t)$ be the state variables at age t that are either endogenous or stochastic. Then the value function at age t can be written as

$$\begin{aligned} V_t[S_t, X_t] &= f_t(m_t, c_{t-1}, d_t, X_t) + 1(m_t > 1) \varepsilon_t - D_{m_{t-1}} 1(m_t = 1) \\ &\quad + \beta E_{c_t, \varepsilon_{t+1}} \left\{ \max_{m_{t+1} \in F(m_t)} (V_{t+1}[S_{t+1}, X_{t+1}]) \mid \hat{\theta}_{d_t}, \varepsilon_t, c_{t-1} \right\} \end{aligned} \quad (3.6)$$

¹⁶A problem with this justification is that it would lead to an unusually large number of separations at \bar{t}_d as couples discover big (negative) mistakes in their beliefs about θ .

¹⁷In fact, given our maximum likelihood estimate of ρ ($= .498$), $\bar{t}_d = 6$ is not long enough for the updating rule for $\hat{\theta}_{d_t}$ to be a good approximation.

where $F(m_t)$ is the feasible set of choices,

$$f_t(m_t, c_{t-1}, d_t, X_t) = \alpha_0(m_t) + c_{t-1}\alpha_c(m_t) + d_t\alpha_d(m_t) + t\alpha_t(m_t) + X_t\alpha_x(m_t) + \mu(m_t) \quad (3.7)$$

is the deterministic utility flow¹⁸ given S_t with $\mu(m_t)$ varying across m_t and across the population (to allow for unobserved heterogeneity),¹⁹ and

$$D_m = X_t\delta_m \quad (3.8)$$

is a separation cost from state m ($D_1 = 0$, $D_2 > 0$, $D_3 > 0$). Note that there are no costs associated with forming a cohabitation or a marriage. Also, note that c_{t-1} affects flows rather than c_t ; this avoids the coherency problem described in Heckman (1979) and Schmidt (1982). Equation (3.7) implies that children are valued as a relationship-specific investment. Note that, in equation (3.3), birth probabilities are allowed to depend upon characteristics of the relationship (e.g., type and duration), but they are not allowed to depend directly upon θ , which would be necessary for treating the investments of this type as completely endogenous.

Value functions are solved by assuming that there is some age t^* such that no decisions are made after t^* and another age t^{**} at the end of which the person dies. This implies that

$$V_{t^*}[S_{t^*}, X_{t^*}] = \sum_{s=t^*}^{t^{**}} \beta^{s-t^*} f_s(m_{t^*}, c_{s-1}, d_s, X_s) + 1(m_{t^*} > 1) E \left[\sum_{t=t^*}^{t^{**}} \beta^{t-t^*} \varepsilon_t \mid \hat{\theta}_{d_{t^*}}, \varepsilon_{t^*} \right] - D_{m_{t^*-1}} 1(m_{t^*} = 1)$$

where $c_{1s} = c_{1t^*}$, $c_{2s} = \min(c_{2t^*} + s - t^*, \bar{n}_{ca})$, $d_s = \min(d_{t^*} + s - t^*, \bar{t}_d)$, and X_s changes in a nonstochastic, exogenous way. Then $V_t[S_t, X_t]$ can be evaluated iteratively for all $t < t^*$.

¹⁸In the empirical work, we allow age t and duration of relationship d_t to affect flow nonlinearly through a spline function.

¹⁹It is theoretically possible that the distribution of θ changes with age because of equilibrium marriage market affects. To some degree, we capture this by allowing flows at t to depend upon age at t . But we miss this equilibrium effect to the degree equilibrium conditions imply that flows should also depend upon the age at the beginning of a relationship and to the degree it implies that the functional form of the distribution should change with age at the beginning of a relationship.

4. Implications of the Model

In this section, we describe behavior of agents in the model. These characteristics of behavior are formalized in the Appendices B, C, and D in Brien, Lillard, and Stern (2000). First, following standard methods, it is straightforward to show that the value functions and their derivatives with respect to ε_t and $\hat{\theta}_t$ are bounded. In particular, let $V_t(m_t, m_{t-1}, \hat{\theta}_t, \varepsilon_t)$ be the value function with other arguments suppressed. Then $\partial V_t(m_t, m_{t-1}, \hat{\theta}_t, \varepsilon_t) / \partial \varepsilon_t > 1$ for $m_t > 1$ because the flow increases by 1 and increases in ε_t result in an increase $\hat{\theta}_t$, implying better flows in the future. On the other hand, $\partial V_t(1, m_{t-1}, \hat{\theta}_t, \varepsilon_t) / \partial \varepsilon_t = 0$ because the agent does not receive utility from ε_t if single and today's ε_t provides no information about ε_{t+1} with a new potential match.

4.1. Reservation Values

Behavior is described by a set of reservation values. Let $\varepsilon_t^*(m_t, 1)$ be a reservation value for single people such that

$$\begin{aligned} V_t(m_t, 1, \hat{\theta}_t, \varepsilon_t) &> V_t(1, 1, \hat{\theta}_t, \varepsilon_t) \quad \forall \varepsilon_t > \varepsilon_t^*(m_t, 1); \\ V_t(m_t, 1, \hat{\theta}_t, \varepsilon_t) &< V_t(1, 1, \hat{\theta}_t, \varepsilon_t) \quad \forall \varepsilon_t < \varepsilon_t^*(m_t, 1). \end{aligned}$$

Such a reservation value exists because $V_t(m_t, 1, \hat{\theta}_t, \varepsilon_t) - V_t(1, 1, \hat{\theta}_t, \varepsilon_t)$ is increasing in ε_t at a rate bounded from below. Also, by a similar argument, there exists a reservation wage $\varepsilon_t^*(1, m_{t-1})$ such that

$$\begin{aligned} V_t(m_{t-1}, m_{t-1}, \hat{\theta}_t, \varepsilon_t) &> V_t(1, m_{t-1}, \hat{\theta}_t, \varepsilon_t) \quad \forall \varepsilon_t > \varepsilon_t^*(1, m_{t-1}); \\ V_t(m_{t-1}, m_{t-1}, \hat{\theta}_t, \varepsilon_t) &< V_t(1, m_{t-1}, \hat{\theta}_t, \varepsilon_t) \quad \forall \varepsilon_t < \varepsilon_t^*(1, m_{t-1}). \end{aligned}$$

These reservation values rank coresidential states relative to being single, but they don't rank marriage relative to cohabitation.

It is difficult to characterize behavior with any more detail without making additional assumptions. In particular, we need to make some assumptions that cause both cohabitation and marriage to exist. We hypothesize that people cohabit to avoid potential divorce costs especially when they are uncertain about the quality of the match. We also hypothesize that they marry due to an additional

utility bonus.²⁰ Specifically, we assume that

$$D_3 > D_2 > 0 \quad (4.1)$$

and

$$f_t(3, c_{t-1}, d_t, X_t) > f_t(2, c_{t-1}, d_t, X_t). \quad (4.2)$$

If the strong inequality in equation (4.1) does not hold, then marriage dominates cohabitation and no one cohabits. If the strong inequality in equation (4.2) does not hold, then cohabitation dominates marriage and no one marries. We assume that both equations (4.1) and (4.2) hold to ensure the existence of both marriage and cohabitation.²¹

Because $D_3 > D_2$, there will be some values of ε_t where a married couple will remain married but a cohabiting couple will choose to separate;

$$\varepsilon_t^*(1, 3) < \varepsilon_t^*(1, 2). \quad (4.3)$$

This implies that

$$\frac{\partial V_t(3, m_{t-1}, \hat{\theta}_t, \varepsilon_t)}{\partial \varepsilon_t} > \frac{\partial V_t(2, m_{t-1}, \hat{\theta}_t, \varepsilon_t)}{\partial \varepsilon_t}.$$

The increase in flow as ε_t increases is the same for cohabitation and marriage. But the expected future value functions are higher for marriage because of the values of ε_t between $\varepsilon_t^*(1, 3)$ and $\varepsilon_t^*(1, 2)$. Thus, a reservation value $\varepsilon_t^{**}(3, m_{t-1})$ determining whether to marry exists:

$$\begin{aligned} V_t(3, m_{t-1}, \hat{\theta}_t, \varepsilon_t) &> \max_{m < 3} V_t(m, m_{t-1}, \hat{\theta}_t, \varepsilon_t) \quad \forall \varepsilon_t > \varepsilon_t^{**}(3, m_{t-1}); \\ V_t(3, m_{t-1}, \hat{\theta}_t, \varepsilon_t) &< \max_{m < 3} V_t(m, m_{t-1}, \hat{\theta}_t, \varepsilon_t) \quad \forall \varepsilon_t < \varepsilon_t^{**}(3, m_{t-1}). \end{aligned}$$

Since $\varepsilon_t^{**}(3, m_{t-1})$ is finite, some people choose to marry. For some values of separation costs and flows, some people choose to cohabit because, when $f_t(3, c_{t-1}, d_t, X_t) = f_t(2, c_{t-1}, d_t, X_t)$, cohabitation dominates marriage and value functions are continuous in flow functions.

²⁰Neither of these assumptions are imposed on the utility flow and the separation cost functions in the estimation procedure discussed below.

²¹We also need to assume that $D_3 - D_2$ grow at a rate over t or d_t no greater than β^{-1} to rule out the possibility of divorce costs rising so rapidly that agents divorce early to avoid being trapped later in relationships.

4.2. Separation Probabilities

The structure of the model allows us to make statements regarding the risk of divorce for marriages preceded by a nonmarital cohabitation and those that were not. As noted above, previous research has found that marriages preceded by a nonmarital cohabitation were less stable. Lillard, Brien, and Waite (1995) empirically show that this difference can be accounted for by a selection effect. Consider a relationship of length $t-1-\tau$ for which the first $k \geq 0$ periods were a nonmarital cohabitation and the last $t-1-\tau-k$ periods were a marriage. Let $\tilde{P}_t(k, \tau) = \Pr [m_t = 1 \mid m_{\tau-1} = 1, m_s = 2 \forall \tau + k > s \geq \tau, m_s = 3 \forall t-1 \geq s \geq \tau + k]$ be the probability of divorce for this type of relationship. The model implies that $\tilde{P}_t(k, \tau)$ increases in k . Couples who began their relationship as a cohabitation had lower values of $\hat{\theta}$ than those who moved straight into marriage. Because $\hat{\theta}$ is positively serially correlated within a relationship, those married couples who had a period of cohabitation have lower values of $\hat{\theta}$ than those married couples who had no period of cohabitation. This is true even after considering the selection associated with cohabiting couples into marriage. Lower values of $\hat{\theta}$ imply higher divorce probabilities. An implication of this result is that, if it were required that all marriages were preceded by a cohabitation, then divorce rates would fall.

One's intuition and the literature might suggest that the model has other properties concerning transition probabilities. Let the transition probability at age t from m_t to m_{t+1} be $P_t[m_{t+1} \mid m_t]$ (with other possible arguments suppressed). Equation (4.3) implies that $P_t[1 \mid 2] > P_t[1 \mid 3]$.

One's intuition might suggest that $\partial P_t[1 \mid 3] / \partial D_3 < 0$. Divorce costs reduce the value of becoming single when married and also reduce the value of becoming married when single. Marriage utility flows increase the value of being married relative to other relationship states. However, nonlinearities in future flows and future transition probabilities complicate the issue so that these effects can not be signed. A sufficient condition for these complications to not occur is that age, duration, and children do not affect utility flows. These are very strong restrictions.

Also, one might think that the reservation value associated with divorce $\varepsilon_t^*(1, 3)$ is increasing in duration because the "option value" associated with learning about a partner is diminishing in duration. This is not the case. A key result in Jovanovic (1979) is that there is a negative duration dependence associated with exiting a job caused by unobserved heterogeneity. However, Jovanovic shows that the negative duration dependence occurs only after some critical point in tenure. Separation rates should decline with duration because of the unobserved

heterogeneity in ε_t and θ (see Jovanovic 1979) and because of potential positive increases in utility flows with duration ($\partial f_t(3, c_{t-1}, d_t, X_t) / \partial d_t$ may be positive). The second effect is consistent with Becker’s notion of increasing relationship-specific capital. In general, we can not show such a result for the duration in any particular state. This is partly due to the age, duration, and children utility flow effects and partly due to the finite horizon assumption.

A special case where we can sign some of these effects is if $f_m = f_t(m, c_{t-1}, d_t, X_t)$ does not depend upon t , c_{t-1} , d_t , or X_t . For this special case, divorce hazards decrease with increasing divorce costs and increasing marriage flows. If we are willing to also assume that $t^* \rightarrow \infty$, then, for large enough t and d_t , $P_t[3 | 3, d_t] > P_{t-1}[3 | 3, d_{t-1}]$; this is the analog to the result in Jovanovic (1979).

5. The NLS72 Data

The parameters of the model are estimated using data from the National Longitudinal Study of the High School Class of 1972 (NLS72). More than 22,000 individuals were first interviewed when they were high school seniors in the spring of 1972 with follow-up interviews in 1973, 1974, 1976, 1979, and, for a limited group, 1986. The 1986 follow-up was unique in that it was an unequal probability sub-sample that included with certainty all individuals in the original sample who had a marital disruption or who were single parents. NLS72 provides information on family background, educational achievement, and a variety of other economic and demographic details from the age the respondents left high school until they reached their early thirties. The female participants of the 1986, or fifth, follow-up constitute the basic sample for the analysis of this paper.²²

The NLS72 data contain detailed histories of cohabitation, marriage, and fertility, asked in the fifth follow-up.²³ This follow-up contains starting and ending

²²The sampling scheme used in the fifth follow-up implies a choice-based sampling problem. We deal with that, following Manski and Lerman (1979), by weighting each observation appropriately.

²³Cohabitation is defined in the NLS72 as living with someone of the opposite sex in a “marriage-like” relationship for at least one month. A concern with retrospective histories is the mis-reporting of these events. Bumpass and Sweet (1989) suggest that the NLS72 may underreport cohabitation events in comparison to the National Survey of Families and Households. Other research suggests reasonably comparable rates between the two surveys (Lillard, Brien, and Waite 1995). With regard to marital histories in general, Lillard and Waite (1989) suggest that, in the Panel Study of Income Dynamics, retrospective marriage histories match contemporaneous reports fairly well.

dates for up to three coresidential relationships, either cohabitations or marriages. The data specifically identify cohabitations followed by marriage to the same partner. Using these histories, we created family formation histories for each individual. These include the dates of the beginning and ending of each cohabitation, dates of marriage, separation, and divorce for multiple events of each type, and dates of conception for each child. We matched these detailed retrospective family histories for individuals to data on their educational achievement and enrollment status, state of residence, and other background information. After excluding observations with bad or missing data, we are left with data for 6,118 white and black women. For the purpose of estimation, we examine the status of individuals every 6 months from the time the respondent is age 16 until the end of the panel, age 32.²⁴ By the end of the panel, approximately 65% of the respondents have had one coresidential partner (either a marriage or cohabitation), 15% have had two, and 2% have had three or more.

6. Estimation

The set of parameters to estimate include γ (the birth equation coefficients in equation (3.3)), α (the utility flow equation coefficients in equation (3.7)), δ (the separation cost equation coefficients in equation (3.8)), ρ (the serial correlation coefficient), and σ_θ^2 (the variance of θ). The value of σ_η is restricted to be 1.649 ($= \exp\{.5\}$) because it is unidentified in the same sense as all discrete choice models having an unidentified variance term. All of the α and δ parameters and σ_θ are measured relative to σ_η . The first step in estimation is constructing likelihood contributions for each sample individual.²⁵ Assume data is observed for $t = 1, 2, \dots, T$, let $\tilde{m} = (m_1, m_2, \dots, m_T)$ be the vector of relationship states selected by the agent, and let $\tilde{b} = (b_1, b_2, \dots, b_T)$ be the vector of birth realizations. The likelihood contribution for the agent is the probability of the path of observed relationship choices and birth realizations.

Let $S_t = (m_{t-1}, c_{t-1}, d_t)$ be the vector of state variables excluding m_t , $\hat{\theta}_{d_t}$, and ε_t .²⁶ First, we need to condition on the value of the unobserved heterogeneity in

²⁴This aggregation implies that we will miss some transitions in and out of relationships. The alternative to the semi-annual observations is to use monthly data. This, however, would be too computationally expensive for our proposed methods. We have examined the severity of this problem, and found that the effects of aggregation were pretty small.

²⁵We will suppress a person index until it is necessary.

²⁶Note that duration, d_{t+1} , is also updated depending on m_t and m_{t+1} .

the flow equations, $\mu = [\mu(2), \mu(3)]$.²⁷ Then the path probability conditional on μ can be written as

$$\Gamma(\mu) = \Pr \left[V_t(m_t, S_t, \hat{\theta}_{d_t}, \varepsilon_t) \geq V_t(j, S_t, \hat{\theta}_{d_t}, \varepsilon_t) \right. \\ \left. \forall j \in F(m_{t-1}); b_t \quad \forall 2 \leq t \leq T \right]. \quad (6.1)$$

Equation (6.1) is the probability that, for each period $2 \leq t \leq T$, the value of the observed choice $V_t(m_t, S_t, \hat{\theta}_{d_t}, \varepsilon_t)$ is at least as large as the value of all choices in the choice set $F(m_{t-1})$ and observing a birth outcome b_t . This can be written as

$$\prod_{t=3}^T \Pr[V_t(m_t, S_t, \hat{\theta}_{d_t}, \varepsilon_t) \geq V_t(j, S_t, \hat{\theta}_{d_t}, \varepsilon_t)] \\ \forall j \in F(m_{t-1}); b_t \mid \varepsilon_{t-1}, \hat{\theta}_{d_{t-1}} \bullet \\ \Pr[V_2(m_2, S_2, \hat{\theta}_2, \varepsilon_2) \geq V_2(j, S_2, \hat{\theta}_2, \varepsilon_2) \forall j \in F(m_1); b_2] \quad (6.2)$$

where $\hat{\theta}_{d_t}$ is a function of ε_t and $\hat{\theta}_{d_{t-1}}$ as shown in Appendix A.

Each of the transition probabilities in equation (6.2) can be written as

$$\Gamma_t(\mu) = \Pr[m_t = k, b_t = b] \\ = p_t^b (1 - p_t)^{1-b} \frac{1}{\sigma_\eta} \int_{-\infty}^{\infty} \phi \left(\frac{\varepsilon_t - [(1 - \rho)\hat{\theta}_{d_{t-1}} + \rho\varepsilon_{t-1}]}{\sigma_\eta} \right) \bullet \\ 1 [V_t(k, S_t, \hat{\theta}_{d_t}, \varepsilon_t) \geq V_t(j, S_t, \hat{\theta}_{d_t}, \varepsilon_t) \forall j \in F(m_{t-1})] d\varepsilon_t. \quad (6.3)$$

The likelihood contribution for an individual is the integral of $\Gamma(\mu)$ integrated over the joint density of μ . The path probabilities are too difficult to evaluate analytically partially because they depend upon unobserved heterogeneity μ (and therefore must be integrated over μ) and partially because the integrals in equation (6.3) have no closed form. However, they can be simulated using a modified GHK algorithm (Geweke 1991, Hajivassiliou 1990, and Keane 1994). The complete algorithm and other related issues are discussed in Appendix A.

²⁷There is also unobserved heterogeneity in $\mu(1)$. But, since we observe only discrete choices, we can identify only μ terms relative to a base choice. Thus, we set $\mu(1) = 0$ and measure $\mu(2)$ and $\mu(3)$ relative to $\mu(1)$.

7. Estimation Results

7.1. Childbearing Decisions

Prior to estimating the full model of relationship choices, following Rust (1988), we obtain parameter estimates for the probability that a woman conceives a child in a particular period. As described above, because of computational costs, we do not model childbearing decisions as part of a utility optimization problem, but we allow the probability that the agent conceives a child to be a function of her characteristics and relationship status. The predicted probabilities are then used in the estimation of the relationship model. Logistic regressions were estimated for each woman and year represented in the full panel of the NLS72 data. The outcome modeled is an indicator variable that equals one if the woman conceived a child in the period and zero otherwise.²⁸ Control variables include the woman’s race (black, white), religion (Catholic, non-Catholic), years of education, school enrollment status, region of residence (south, midwest, west, northeast), and piecewise linear splines in age, duration of current relationship (including years single), number of children, and the age of the youngest child.²⁹

The results of this estimation, which are presented in Table 1, suggest a statistically significant role of many of the covariates and that their impact varies across relationship status. For example, we find that black women are much more likely than white women to have a child either outside of a coresidential relationship or in a nonmarital cohabitation. This is consistent with the literature documenting the large racial differences in nonmarital childbearing (e.g., U.S. Dept. of Health and Human Services 1995). We also find that Catholics are less likely to have a child while single and more likely to have a child while married. More education has a negative effect on the probability of having a child while single or cohabiting, and a positive but insignificant effect while married. Being enrolled in school leads to a lower probability of childbearing regardless of the woman’s relationship status. Finally, a woman’s fertility history, including both the number and age of children, plays an important role for all three relationship statuses.

It is difficult to see directly from the table how the probability of conceiving varies with relationship type. In fact, the parameter estimates imply that, for over 99% of the observations, the probability of conceiving is higher when cohabiting or

²⁸The data includes only the birthdates for children who were carried to term. The conception is dated such that it falls in the period encompassing nine months prior to the birthdate.

²⁹Number of children and age of youngest child are subject to the upper bounds discussed in the theoretical model.

married than when single, and for 92% of the sample observations, the probability of conceiving is higher when married than when cohabiting. This is consistent with our notions of the effect of relationship type on fertility.

7.2. Relationship Choices

7.2.1. Parameter Estimates

The results for the model of relationship choices are presented in Table 2. This table is composed of three panels. Panel A contains the estimated utility flow coefficients. Panel B presents the estimates of the costs of separating from a coresidential relationship. Finally, Panel C presents estimates of two auxiliary parameters; the standard deviation of the match quality θ and the autocorrelation term ρ . It should be noted that the estimates in all three panels are statistically significant at the 5% level.³⁰

First consider Panel A and the factors that affect an individual’s utility flow. The first column presents the estimates for cohabitation and the second column is for marriage. Both sets of estimates should be interpreted as the incremental utility flow associated with the particular state (relative to being single) when the associated covariate increases by one unit. For example, the cohabitation coefficient on Catholic, -0.288 , means that, holding all else constant, Catholics receive a utility flow 0.288 units less than non-Catholics (relative to the utility flow received when single). In this model, utility units are measured relative to the standard deviation of η , the innovation to utility flows. Recall that the value of σ_η is restricted to be 1.649 ($= \exp\{.5\}$). Thus, for example, the coefficient on Catholic means that a Catholic’s cohabitation utility flow is $0.288/1.649 = .175$ standard deviations of η less than a non-Catholic’s utility flow.

The model contains a number of variables similar to those used in the child-bearing equations described above. These include indicator variables for whether the respondent is black, is Catholic, is enrolled in school, and geographic region of residence. Years of education, the number of children, and the age of the youngest child are included as ordered integer variables. The respondent’s age and duration in a relationship status are measured with piecewise linear spline variables. The age, duration, and child age variable are all measured in 6 month time periods.

³⁰Frequently, t-statistics in nonlinear structural models are very large. See, for example, Pakes (1986), Rust (1987), or Berry (1992).

Consistent with other literature on marriage and marriage markets, we find that black women receive less utility than white women in a coresidential relationship. The effect is larger for marriage than for cohabitation. We also find that being Catholic is associated with having a lower utility in cohabitation and marriage. It is important to note that our utility effect is not due to a higher potential divorce cost for Catholics. As will be discussed below, the Catholic prohibition on divorce is captured in Panel B.³¹ The pattern of the region dummy variables imply that, relative to the other regions, individuals in the south gain the most from being in a relationship and individuals in the west gain the least.

More education implies a lower level of utility in either relationship state. As suggested above, this may occur because of greater employment opportunities for more highly educated women outside of a relationship. Similarly, being enrolled in school leads to less utility when in a coresidential relationship. This effect is slightly larger for marriage.

Finally, the utility associated with being in a relationship increases in the number of children for cohabitations and decreases for marriages. While this result may seem somewhat counterintuitive (since individuals have a higher probability of conceiving a child when in a marriage), there are aspects of the relationship between childbearing and marriage that make this possible. Specifically, the presence of children could influence both the transition into and out of marriage. The negative coefficient for marriages, which is relatively small, is caused by the sample hazard rate into marriage by women decreasing with the number of children born prior to marriage, a result found elsewhere in the demographic literature. Short of endogenizing the childbearing decision, a possible improvement in our empirical implementation might be to distinguish between children born prior to the marriage and children born during the marriage. The negative utility coefficient may in part be due the inability of the data to identify whether the marriage partner is the biological father of the woman's children. Altering the model to make this distinction is problematic, however, because it significantly increases the size of the state space and because it is very weakly identified by the data. In particular, sample hazard rates into marriage can not help identify the effect of children born during the marriage.³² It should be further noted that

³¹It is important to note that Catholic might be capturing social or ethnic differences rather than religious differences.

³²Technically, sample hazard rates into marriage provide some information about the effect of number of children on marriage utility but only because birth rates vary by relationship types. Identification through this mechanism relies on functional form assumptions about how future

the negative effect for marriage is not due to the relationship between the hazard out of marriage and the number of children. This effect is captured in the divorce costs coefficients discussed below.

This analysis also begs the question of why the coefficient for children is positive for cohabitation. It may be that relationships are useful in raising children in the sense of Becker's model of household behavior. If one marries, one gains the benefit of having another adult but loses any payments (e.g. child support, AFDC) associated with not being married. While cohabitation may affect AFDC eligibility (Moffitt, Reville, and Winkler 1998), the negative part of the tradeoff does not exist (or at least is not as important) for cohabiting. The results also indicate that the utility value of children declines as the children age.

The change in the utility flow over the course of a relationship is captured by the duration coefficients. Since the cohabitation duration effects are larger than those found for marriage, the utility flow premium associated with marriage dissipates as the relationship progresses. Consider a 20 year old, white, non-Catholic woman with 12 years of schooling, living in the south with no children and not enrolled in school. At the beginning of the relationship, she has a marriage utility flow premium of 1.054, and, after 3 years of a relationship, she has a marriage utility flow premium of 0.816. The negative duration dependence in the divorce rate exhibited in the data has two sources: the positive coefficients on duration in Table 2 and the existence of unobserved heterogeneity in values of the match quality θ caused by variation in the unobserved component μ and randomness in θ conditional on μ .

Panel B considers the utility costs associated with separating from a relationship. The control variables are the number of children, the age of the youngest child, and an indicator for whether the respondent is Catholic. We find higher divorce costs for Catholics. Interestingly, they also seem to pay a higher cost of separating from a nonmarital cohabitation. As expected, the presence of children raise the cost of separating from a relationship, and the cost is higher for exiting a marriage than a cohabitation. Regardless of religion and child characteristics, separation costs are always higher for marriage than for cohabitation.³³ This result, which was not imposed on the parameters by the estimation procedure, reflects, among other things, the legal costs of ending a formal marriage and religious prohibitions on divorce.

utility depends on present choices; thus it is weak.

³³This is true because the number of children is censored at 4 and age of the youngest child is censored at 2 years of age.

Panel C provides estimates of σ_θ , σ_μ , and ρ . Note that σ_θ is 0.59 which is significantly smaller than σ_η suggesting that there is more variance across time within a relationship than across potential relationships. The estimate of σ_μ is even smaller, suggesting that most of the variation in match values is specific to the match and not specific to the agent across matches. However, allowing for variation in μ is important to explaining the data. The estimate of the serial correlation term ρ is 0.498, suggesting moderate positive serial correlation in η shocks. Given our estimate of ρ , the $Var(\theta | \hat{\theta}_{d_t})$ six periods into a relationship is 0.28. In contrast, if we assume that there is no serial correlation ($\rho = 0$), we observe a sharper decline in the variance, implying faster learning; six periods into a relationship, the $Var(\theta | \hat{\theta}_{d_t})$ is 0.20.

Overall, the parameter estimates have interesting implications for the effect of the size of relative utility flows on relationship choices. Consider, for example, the hypothetical 20 year old, white woman described above. As noted above, for such a person, the average differential between the marriage utility flow and cohabitation utility flow for starting a new relationship is 1.054. Note that there is variation in utility flows due to the randomness of θ and η . In fact, the marriage utility flow differential (1.054) is small relative to the standard deviation of the random components. This suggests that even small marriage utility flow premiums (relative to cohabitation utility flows) can generate high marriage rates even with large differences in separation costs as long as the relationship utility flows associated with being in a relationship are large relative to the utility flow associated with being single. In such a case, a person is willing to incur the risk associated with divorce to enjoy a modest marriage utility flow premium because divorce is an unlikely event.

Another implication of the estimates is that it is impossible to generate a hypothetical person whose utility flow to cohabitation exceeds the utility flow to marriage at the beginning of a relationship. There was nothing in the estimation procedure that required this to happen. This result supports the theoretical model assuming a positive premium to marriage (relative to cohabitation) as the cause of couples marrying despite high divorce costs. As the model shows, couples still sometimes cohabit because of higher separation costs for marriage than cohabitation and because utility flows change over the course of the relationship.

7.3. Specification Tests

Using our estimates, we simulate a sample of 40,000 individuals and then compare their behavior to that observed in the data sample. Results of this exercise are displayed in Figures 2 and 3 and Table 3. Figure 2 displays the simulated and data sample proportions of women in each of the three states as a function of age. We predict all three proportions extremely precisely at all ages. Table 3 reports censored χ^2 goodness-of-fit tests for different subsamples of our data disaggregating by race (black and white), religion (Catholic and not), education (12, 14, and 16 years of education), region (4 regions), and age (18, 22, 26, and 30 years old). We censor the χ^2_1 statistics associated with each cell at the 1% right tail to mitigate the undue influence of large outliers. The results of these goodness-of-fit tests also suggest we are fitting the data very well for transitions from marriage but, somewhere, the model is misspecified with respect to transitions from single and cohabiting. The results are similar if we disaggregate by race, religion, and age. One should note that, in most specification tests of structural discrete choice models, formal specification tests reject the model (see, for example, Berkovec and Stern 1991, Keane and Wolpin 1997) or use goodness-of-fit measures based on a MSE criterion which is less formal and less sensitive to small denominators (see, for example, Pakes 1986, Berry 1992).

Figure 3 compares simulated and data sample hazard rates. The hazard rates by age for each transition are presented in the separate panels of Figure 3. Panel A shows simulated and data sample hazard rates from single to cohabitation. Combined with the information in Panel C, the hazard rates from cohabitation to single, it appears that we are overpredicting transitions between cohabitation and single in both directions and by similar amounts.

Panel B shows simulated and data sample hazard rates from single to married. We overpredict marriage rates at ages after age 25, but, surprisingly, it does not show up in the deviations between predicted and sample proportions in marriage in Figure 2.

Overall, our specification tests suggest that we are fitting the data very well, especially in terms of proportions in each state. There are a few areas where we overpredict transitions.

We also tested for different forms of heteroskedasticity using Lagrange Multiplier tests similar to those described in Breusch and Pagan (1979). In particular, we allowed ρ , σ_θ^2 , and σ_μ^2 to vary with race, religion, education, and region of residence. Table 4 provides support for heteroskedasticity in σ_θ^2 but not in ρ or σ_μ^2 . The size of σ_θ^2 increases with Black and Catholic and decreases with years

of education. The results for Blacks and Catholics might be a sign of assortive mating. The pool of mates for blacks and Catholics is smaller than for whites and non-Catholics. Whites and non-Catholics may be able to specialize their search better, thus leading to smaller variances. This argument does not work as well for education.

Finally, we compared (weighted) residuals across personal characteristics that would be difficult to incorporate in a formal Lagrange Multiplier test because it would involve adding a state variable to the model thus making estimation or testing infeasible. As a first test, we aggregated residuals across states of residence conditional on race, religion, age group, and region to check for the existence of other state effects. If we construct a quadratic form in such residuals weighted by their inverse covariance matrix and approximate its distribution with χ_q^2 where q is the number of combinations of states of residence and relationship states m for which we have residuals,³⁴ then we reject the null hypothesis that there are no other state effects.³⁵ However, if we regress the aggregated residuals (based on transitions from single to marriage and from marriage to single) on the three divorce cost variables used in Friedberg (1998),³⁶ we find no significant estimated effects of the divorce laws and overall R^2 less than 0.1%. Thus, while our results suggest that there may be other state effects, they are not captured by variation in divorce laws. This result is somewhat different than that reported by Friedberg (1998) who found that, after controlling for unobserved differences across states, changes in divorce laws played a significant role in explaining the time-series variation in divorce rates.

A second test of this “pseudo-Lagrange Multiplier” type compares residuals between those who have not finished their first relationship and those who have. More precisely, let

$$e_{it} = 1(\exists s < s' < t : m_{is} > 1, m_{is'} = 1)$$

be an indicator for separation from a relationship. We want to test whether e_{it} is correlated with the residuals. If we construct a quadratic form in such residuals weighted by their inverse covariance matrix and approximate its distribution with χ_q^2 , then we reject the null hypothesis that behavior between first

³⁴We limited our observations to those cells where there were at least four observations.

³⁵The overall χ^2 statistic is 260.5 with 82 degrees of freedom.

³⁶The three variables are dummies for a) whether the state allowed unilateral divorce, b) whether there is a waiting period before a divorce can occur, and c) whether settlements are no-fault. See Friedberg (1998) for a complete description of these variables.

and subsequent relationships is the same.³⁷ However, the difference in average residuals is small: for transitions originating from being single, the difference in average residuals for $e_{it} = 1$ and $e_{it} = 0$ for transitions into marriage is 0.018; for transitions originating from being married, the difference in average residuals for $e_{it} = 1$ and $e_{it} = 0$ for transitions into single is 0.011. Thus, our results suggest there is little to be gained by adding a state variable for the existence of prior relationships. In the prior literature, however, a number of researchers have found important differences between first and subsequent relationships. Bumpass and Sweet (1989), for example, report higher rates of cohabitation for those who had experienced a prior relationship. Upchurch, Lillard and Panis (1999) use a hazard framework to examine the determinants of the waiting time to marriage and marital dissolution and find that, when not accounting for unobserved differences across individuals, prior marriages significantly increases the hazard of both events. When estimating the processes jointly, including a fertility decision, they find this result disappears and, in the case of the waiting time to marriage, prior marriages significantly decrease the hazard of marriage.

7.4. Comparative Dynamics

In this section, we conduct a number of experiments to evaluate potential family policies. The first experiment is a lowering of the divorce cost. This type of experiment could help evaluate the changes in family structure that may have occurred during the 1970's and 1980's when there was a substantial relaxation in divorce laws (Friedberg 1998) and significant changes in the marriage penalty associated with the U.S. tax code (Whittington and Alm 1997). Mechanically, this exercise is accomplished by decreasing the parameter estimate on the constant term in the divorce cost equation by 0.1. To help gauge in monetary terms a parameter change of this size, we rely on two papers that relate marital status transitions to variables denoted in dollars. Whittington and Alm (1997) find that the effect of a \$1000 increase in the marriage penalty on the annual divorce rate is 0.004 (Table 2, p 399) for the average person. This translates into an effect of 0.105 utils where a util is the unit of measure in our model.³⁸ In a related paper, Alm and Whittington (1999) find that the effect of a \$1000 increase in the marriage penalty on the annual marriage rate is -0.002 (Table 3) for the average person. This translates into an effect of 0.027 utils. Note that our model and most other

³⁷The χ^2 statistic is 525.7 with 172 degrees of freedom.

³⁸See Appendix B for an explanation of this transformation.

economic models of marriage and divorce suggest that these two numbers should be the same magnitude. Thus, for our analysis, we take a geometric average of these and assume \$1000 is worth 0.05 utils; our experiment corresponds to a \$2000 decrease in divorce costs. We should be clear about how we are combining the Alm and Whittington results with the parameter estimates in our paper. Alm and Whittington is really providing the overall magnitude of the effect. Our contribution is to translate that effect into this model so that the specific policy question they addressed can be compared to other policy questions of relevance that they could not have analyzed given their empirical approach. In effect, we are using the Alm and Whittington results to calibrate one parameter that is unidentified given our data and then performing policy analysis conditional on the calibrated parameter.

There are two ways of analyzing the effects of this experiment: evaluating changes in proportions of the population in each state conditional on age or evaluating hazard rates conditional on age. The proportions of people in each state in the base case are shown in Figure 2. Decreasing divorce costs cause essentially no change in the proportion choosing to be single at any age. Consistent with one's expectations, the experiment results in a decrease in the proportion of individuals choosing to reside in a nonmarital cohabitation and a slight increase in the proportion of the sample that chooses to be married. Overall, the impact of this policy change appears to be in the anticipated direction but fairly small in size. The corresponding hazard rates are presented in the five panels of Figure 4. For example, the top left panel (Panel A) depicts the hazard out of singlehood into cohabitation. There are separate lines for each policy scenario (base vs. decreased divorce cost). The results for leaving the single state (Panels A and B) reveal a slightly higher hazard into marriage and a slightly lower hazard into cohabitation when there is a decrease in the divorce cost. The middle two panels show the hazard out of cohabitation in which we see a slight increase in the hazard for those terminating the relationship (exiting into single). The bottom panel shows the expected, but very small, increase in the hazard of divorce. The small effects from this policy simulation are consistent with the lack of statistical significance of Friedberg's divorce law variables in explaining our residuals reported above. This consistency is especially important and adds credence to the results because, while the magnitude of the policy effect reported here depends upon the estimates in Alm and Whittington's work, the specification test results using Friedberg's data does not.

The second experiment we consider is one in which there is an increase in the

incentives to marry. This might include a decrease in the “marriage penalty” or, more simply, some sort of marriage credit. We operationalize this experiment by increasing the constant term in the marriage equation by 0.1; this is again equivalent to a credit of \$2000 each six months. In this experiment, more people choose to marry, and fewer people choose to cohabit. Figure 5 shows the corresponding hazard rates that are relevant for each of the relationship states. The top two panels show, for single individuals, increases at all ages in the hazard of marriage and decreases in the hazard of cohabitation. Interestingly, and perhaps counter to one’s expectations, the middle two panels show most clearly an increase in the hazard out of cohabitation into singlehood. This may be, in part, due to the relatively small number of individuals predicted to be cohabiting under this policy scenario. Finally, the bottom panel provides fairly mixed results on the hazard of divorce for this experiment.

The final experiment relates to the value of having children. This experiment increases the utility associated with having a child by 0.05 regardless of their relationship status. Such an increase could occur by increasing the exemption on children in the tax code by \$1000 each six months. The results for this policy experiment are remarkably similar to the marriage credit experiment described above. Like we had observed in the marriage credit, the child credit results in a large movement out of cohabitation and into marriage. One possible explanation for the similarity between the experiments may lie in how we treat the childbearing process in the model. Based on the estimates presented in Table 1, women have a significantly higher probability of having a child when married. The presence of the child credit may induce individuals to choose marriage and the accompanying higher probability of having a child.³⁹

8. Conclusion

The theoretical model presented in this paper offers a coherent structure in which to analyze relationship choices. It provides, among other things, an explanation for why individuals choose to live together outside of a formal marriage, why these types of relationships are shorter-lived, and why many nonmarital cohabiting relationships are converted into marriages. While the information-gathering framework undoubtedly only captures part of the reasons individuals choose to form and dissolve relationships, the model appears to satisfactorily replicate many

³⁹We also performed an experiment limiting the utility increase for children to married couples and found no significant difference in behavior.

of the stylized facts surrounding cohabitation and marriage decisions. The model also helps us to interpret other empirical findings such as that marriages beginning as cohabitations are less stable than those that do not. The structural parameters of the model were estimated and have allowed us to further gauge the suitability of the framework and to conduct a variety of experiments.

This work suggests a number of avenues for future research. One obvious shortcoming of the present framework is the manner in which childbearing is not treated as a fully endogenous decision. As discussed above, the various behaviors were modeled in this way primarily because of the computation costs associated with expanding the state space in an already complex model. The decision to have a child, however, is clearly intertwined with cohabitation and marriage decisions (Brien, Lillard, and Waite 1999) and advances in the numerical methods used to solve and estimate dynamic models are needed to make this problem tractable. Another issue is the role of unobservables. One might consider a somewhat different model, not relying on information gathering, where the quality of the match, θ , changes over time. It is worthwhile to derive the implications of such a model and to determine if it is identified from the model we have presented here.

A. Appendix: Estimation Issues

A.1. Simulation of Path Probabilities

The path probability defined in equation (6.3) can be simulated. The simulation algorithm is:

- a) Compute $\Pr [m_2, b_2]$ using equation (6.3);
- b) Draw $\varepsilon_2 \mid m_2$;
- c) Compute $\hat{\theta}_2$ as function of ε_2 ;
- d) Set $t = 2$ and $\Gamma = \Pr [m_2, b_2]$;
- e) Set $t = t + 1$;
- f) Compute $\Pr [m_t, b_t \mid \varepsilon_{t-1}, \hat{\theta}_{d_{t-1}}]$;
- g) Set $\Gamma \leftarrow \Gamma * \Pr [m_t, b_t \mid \varepsilon_{t-1}, \hat{\theta}_{d_{t-1}}]$;
- h) Stop if $t = T$;
- i) Draw $\varepsilon_t \mid m_t, \varepsilon_{t-1}, \hat{\theta}_{d_{t-1}}$;
- j) Compute $\hat{\theta}_{d_t}$ as a function of ε_t and $\hat{\theta}_{d_{t-1}}$;
- k) Go to (e).

A.2. Adjustment for Rare Events

Evaluating equation (6.3) can be done in two parts, $p_t^b(1-p_t)^{1-b}$ and everything else. The first part is straightforward. The second is very complicated; see Appendix E in Brien, Lillard, and Stern (2000) for details. Unfortunately, because of the ordered nature of the relationship choices, most (if not all) values of the parameter vector have the property that there are at least two observations, i and j , where observation i includes a period of cohabitation with zero probability of occurring and j includes a period of marriage with zero probability of occurring. The two conditions together suggest that it would be difficult to find a parameter vector where the likelihood function was positive. One way to fix this problem is to replace zero probability events by small probabilities with informative derivatives. In particular, replace $\Pr [m_t = k, b_t = b]$ in equation (6.3) with

$$p_t^b(1-p_t)^{1-b}[(1-\kappa)\Gamma_t + \kappa Q_t] \quad (\text{A.1})$$

where

$$\begin{aligned} \Gamma_t &= \frac{1}{\sigma_\eta} \int_{-\infty}^{\infty} \phi \left(\frac{\varepsilon_t - [(1-\rho)\hat{\theta}_{d_{t-1}} + \rho\varepsilon_{t-1}]}{\sigma_\eta} \right) \bullet \\ &1 \left[V_t(k, S_t, \hat{\theta}_{d_t}, \varepsilon_t) \geq V_t(j, S_t, \hat{\theta}_{d_t}, \varepsilon_t) \forall j \in F(m_{t-1}) \right] d\varepsilon_t, \end{aligned} \quad (\text{A.2})$$

Q_t is a probability function such that $0 < Q_t$, and κ is a small number. Note that if $\kappa = 0$, equation (A.1) reduces to equation (6.3). Thus, equation (A.1) can be thought of as an approximation to equation (6.3) that avoids the problems associated with equation (6.3) equalling zero. A similar idea is used in McFadden (1989). In particular, Q_t should be chosen so that $0 < Q_t$ and so that the partial derivative of Q_t with respect to the parameter vector is informative about the direction one should change the parameter vector to increase Γ_t . A multinomial logit-type probability can be chosen to satisfy these conditions. In particular, let

$$\begin{aligned} Q_t &= \frac{1}{\sigma_\eta} \int_{-\infty}^{\infty} \frac{\exp \{V_t(j, S_t, \hat{\theta}_{d_t}, \varepsilon_t)\}}{\sum_{k \in F(m_{t-1})} \exp \{V_t(k, S_t, \hat{\theta}_{d_t}, \varepsilon_t)\}} \bullet \\ &\phi \left(\frac{\varepsilon_t - [(1-\rho)\hat{\theta}_{d_{t-1}} + \rho\varepsilon_{t-1}]}{\sigma_\eta} \right) d\varepsilon_t \end{aligned} \quad (\text{A.3})$$

for an observation where choice j is chosen at age t . Note that $\partial Q_t / \partial V_t(j, S_t^{b_t^*}, \hat{\theta}_{d_t}, \varepsilon_t) > 0$ and $\partial Q_t / \partial V_t(k, S_t, \hat{\theta}_{d_t}, \varepsilon_t) < 0$ for all $k \in F(m_{t-1}) \neq j$ as desired. If κ is

fixed at a very small number, then it can be argued that a) this is just a small deviation from the model or alternatively b) there really is a small extreme value error in the model but that its size is so small that it has no appreciable effect on value functions.⁴⁰

Let $\Lambda = (\gamma, \alpha, \delta, \rho, \sigma_\theta^2)$ be the parameter vector. Then the maximum simulated likelihood (MSL) estimator of Λ maximizes $\sum_i \log \Gamma_i$ where Γ_i is the path probability for sample person i defined in equation (6.1).

A.3. Numerical Methods for Evaluating the Likelihood Function

In our application, it is too computationally expensive to evaluate the value function for all 6,118 observations, especially with the need to simulate. Some papers in the literature (e.g. Pakes 1986, Stern 1989, Eckstein and Wolpin 1995) attempt to limit the number of times that value functions need to be evaluated by limiting their analysis to homogeneous samples. This is done either by restricting the sample or making severe assumptions about what variables matter. We achieve similar results without relying on these methods. Rather, we create N representative people with different values of X . In our case, X includes race (2 values), religion (2 values), and region (4 values), and we discretize continuous variables $\mu(2)$, $\mu(3)$, and education with 3 values each; thus there are $N = 2 \cdot 2 \cdot 4 \cdot 3 \cdot 3 \cdot 3 = 432$ representative people. We compute the value functions only for these representative agents and “conditional” path probabilities for each sample person i using the value functions for each representative person j “close” to i . Then we use weighted averages of the conditional path probabilities to compute path probabilities for each sample person where the weights are inversely proportional to the distance between the sample person and the representative person.

In particular, let ω_{ij} be the weight associated with representative person j and sample person i where weights are normed so that $\sum_j \omega_{ij} = 1$. Let X^c be the continuous elements of X over which we need to interpolate. Let X^d be the discrete elements of X where, since all relevant values are used to stratify the data (e.g., race, religion, region), there is no need to interpolate. Let $\tilde{X} = (\tilde{X}^c, \tilde{X}^d)$ be the values of $X = (X^c, X^d)$ corresponding to representative people. In our case, the continuous variables X^c are μ and education, denoted X_e . Let $\tilde{\mu}(2)_1$ be the largest “representative” value of $\mu(2)$ less than $\mu(2)_i$ and $\tilde{\mu}(2)_2$ be the smallest

⁴⁰Once the extreme value error is added, the GHK algorithm above needs to be modified. See Brien, Lillard, and Stern (1998).

“representative” value of $\mu(2)$ greater than $\mu(2)_i$, define similar terms for $\mu(3)_i$, and define \widetilde{X}_{e1} and \widetilde{X}_{e2} analogously for education. Then the weighting function (before norming) is given by

$$\omega_{ij} = \prod_{m=2}^3 |\mu(m)_i - \tilde{\mu}(m)_k|^2 |X_{ei} - \widetilde{X}_{ek'}|^2 1(X^d = \widetilde{X}^d) \quad (\text{A.4})$$

where $\tilde{\mu}(m)_k = \tilde{\mu}(m)_1$ if representative person j has $\mu(m) = \tilde{\mu}(2)_2$ and $\tilde{\mu}(m)_k = \tilde{\mu}(2)_2$ if representative person j has $\mu(m) = \tilde{\mu}(m)_1$; $\widetilde{X}_{ek'}$ is defined analogously for the two adjacent representative values of education.⁴¹

The weighting scheme is illustrated in Figure 6 (assuming $X_{ei} - \widetilde{X}_{ek'}$). Consider an observation denoted point A . Let the adjacent vertices for the continuous stratifying variables be denoted by I , II , III , and IV , and let the distances between A and each of the continuous stratifying variables be denoted by z_1 through z_4 . Then the weight for this observation and the representative person corresponding to point I is

$$\omega_{iI} = z_3^2 z_2^2 / [z_3^2 z_2^2 + z_1^2 z_4^2].$$

An alternative to using the weighting scheme described above is to treat μ and education as discrete variables. There is a large literature advocating the use of multi-point discrete distributions for unobserved heterogeneity (see, for example, Mroz 1999). However, there are a number of reasons that treating μ and education as continuous variables is better in this case. First, in the data, education takes on more values than we would like to allow for among representative people. Second, we would have to limit ourselves to distributions with very few mass points (3 or less) to avoid increased computation costs. Third, our own experience suggests that estimating multi-point discrete distributions for unobserved heterogeneity does not behave as well as estimating smoother distributions in derivative based optimization routines. Given the cost of estimation, it is imperative to find a specification that performs well. Finally, we feel it is worth demonstrating the value of this approach so other researchers can use it when they have continuous state variables.

⁴¹The same weighting function is used for interpolating value functions. The weighted average is continuous at interpolation points, and $\partial\omega/\partial x$ exists for any x but is zero at interpolation points. Also, if a sample person has the exact characteristics of a representative person, then only that representative person will receive positive weight.

The path probability for sample person i can be simulated using the modified GHK algorithm described above conditional on the value functions for representative person j . Thus the path probability for sample person i depends on representative person j through j 's value functions, and it depends on sample person i through i 's observed choices and birth realizations. The evaluation of the likelihood function uses the following algorithm (Representative Agent Algorithm):

- 1) Initialize $\Gamma_i = 0$ for all sample people i .
- 2) For each representative person j ,
 - a) evaluate the value function,
 - b) for each sample person i such that $\omega_{ij} > 0$, simulate Γ_{ij} , the path probability for sample person i conditional on representative person j 's value function, and increment Γ_i by $\omega_{ij}\Gamma_{ij}$.

This algorithm increases computation speed by an order of magnitude because the value functions, which are very expensive to evaluate, need to be evaluated only 432 times for each guess of the parameters. If all of the stratifying variables used to define representative people were discrete, then this approach would be exact. However, since at least μ and, arguably, education are continuous variables, one must think of this method as an approximation. If one allows the distance with respect to continuous stratifying variables between representative people to approach zero asymptotically (at a slow rate), then approximation bias disappears asymptotically.⁴²

In the process of evaluating (actually approximating) the value function for a representative person, we also interpolate, using a method similar to Stinebrickner (1996). This approach can be compared with the interpolation methods of Keane and Wolpin (1994). Stinebrickner (1996) is a local regression approximation while Keane and Wolpin (1994) is global regression approximation. The difference between our method and Stinebrickner's is that our weighting function is different. But we also use interpolation in the Representative Agent Algorithm above which is not the case for either Keane and Wolpin (1994) or Stinebrickner (1996).

⁴²Stephen Donald suggested that we could use this algorithm more efficiently by aggregating continuous (and possibly all) exogenous covariates into a single linear index. We could use a relatively fine mesh size and still reduce the number of representative people while increasing the number of covariates we can use as controls. Our only concerns about his suggestion is that, in this problem, one would need two linear indices to capture the effect of covariates on flows and birth probabilities and that the mesh design would depend upon the estimated coefficients in the flow and birth probability equations. Nevertheless, we think this is an excellent area for future investigation.

A.4. Frequency of Rare Events

Given our estimates, it is worthwhile exploring the frequency of such rare events. There are at least three ways to define a rare event. We can focus on a particular choice made in one period and condition on a specific value of μ and a particular representative person (as described above). Then a rare event occurs when it is necessary to use the correction described in equation (A.3) in order to have a positive transition probability $\Gamma_{jt}(\mu)$ defined in equation (6.3) where j indexes the representative person used to evaluate $\Gamma_{jt}(\mu)$. The first panel of Table 5 shows that, at the maximum likelihood estimates, this occurs rarely for single and married periods and relatively frequently for cohabiting periods. However, the results in the first panel seriously overstate the frequency of rare events because the reported rare events are conditioned on values of μ and representative people that do not necessarily correspond to the relevant sample person. Specifically, while particular values of μ for a particular sample person may generate a rare event, there may be other values of μ for that same person that do not generate a rare event. Thus, as an alternative, we might want to know the frequency of rare single period events once we integrate over μ and representative people,

$$\sum_j \omega_j \int \Gamma_{jt}(\mu) d\mu$$

where ω_j is the weight associated with representative person j defined in equation (A.4). The second panel of Table 5 shows that this almost never occurs for single and married periods and only 0.54% of the time for cohabiting periods. Thus, rare events in the data defined this way (i.e., integrated over μ and representative people) are quite rare. Finally, one might want to look at the frequency of “rare paths,”

$$\sum_j \omega_j \int \Gamma_j(\mu) d\mu$$

where $\Gamma_j(\mu)$ is defined in equation (6.1). A rare path would be a path which includes a rare event for each value of μ and representative people so that, even after integration, the probability of the path with the adjustment described in equation (A.3) still would be zero. The third panel of Table 5 shows that this occurs for only 1.16% of the sample paths in the data.

B. Appendix: Transforming Whittington and Alm Estimates into Utils

This appendix shows how to calibrate our utils into dollars using the estimates from Whittington and Alm (1997) and Alm and Whittington (1999) on the effects of the marriage penalty on divorce and marriage rates, respectively. Whittington and Alm (1997) have a standard logit equation which we represent as

$$y^* = X\beta + u \quad (\text{B.1})$$

where u is logistically distributed. They report in their Table 2 results in terms of partial derivatives of probabilities with respect to explanatory variables, $\bar{P}(1 - \bar{P})\hat{\beta}$, where \bar{P} is the mean of the dependent variable. The effect of a \$1000 increase in the marriage penalty on the probability of divorce is 0.004. Given that $\bar{P} = 0.04$, this implies a logit coefficient of $\hat{\beta} = .104$. Whittington and Alm's (1997) time unit is a year, while ours is a half-year. So, assuming equation (B.1) applies to half-year data, define the probability of divorce in two half-years as

$$1 - (1 - P_d)^2 = \left[\frac{\exp\{X\beta\}}{1 + \exp\{X\beta\}} \right]^2.$$

Then, generalizing the approximation in Amemiya (1981),

$$\frac{\partial [1 - (1 - P_d)^2]}{\partial X} = 2(1 - P_d)^2 P_d \beta,$$

which implies

$$\hat{\beta} = \left[2\bar{P}_d (1 - \bar{P}_d)^2 \right]^{-1} \frac{\partial [1 - (1 - P_d)^2]}{\partial X}$$

where $\partial [1 - (1 - P_d)^2] / \partial X$ is Whittington and Alm's reported estimate of 0.004. Also, since $1 - (1 - \bar{P}_d)^2 = 0.04$, $\bar{P}_d = 0.02$. Thus,

$$\hat{\beta} = \left[2(.02)(.98)^2 \right]^{-1} (.004) = 0.1052.$$

Note that changing the time unit has very little effect on $\hat{\beta}$.

Using the same approach to Alm and Whittington (1999) where $\bar{P} = 0.08$, we find

$$\hat{\beta} = \left[2(.04)(.96)^2 \right]^{-1} (-.002) = -0.027.$$

C. References

References

- [1] Akerlof, George. A., Janet L. Yellen, and Michael L. Katz (1996). "An Analysis of Out-of-Wedlock Childbearing in the United States." *Quarterly Journal of Economics*. 61(2): 277-317.
- [2] Alm, James and Leslie Whittington (1999). "For Love or Money? The Impact of Income Taxes on Marriage." *Economica*. forthcoming.
- [3] Amemiya, Takeshi (1981). "Qualitative Response Models: A Survey." *Journal of Economic Literature*. 19(4): 1483-1536.
- [4] Becker, Gary S. (1973). "A Theory of Marriage: Part I." *Journal of Political Economy*. 81(4): 813-846.
- [5] Becker, Gary S. (1991). *A Treatise on the Family*. Cambridge, MA: Harvard University Press.
- [6] Becker, Gary S., Elisabeth M. Landes, and Robert T. Michael (1977). "An Economic Analysis of Marital Instability." *Journal of Political Economy*. 85(6): 1141-1187.
- [7] Bennett, Neil G., A. K. Blanc, and David E. Bloom (1988). "Commitment and the Modern Union: Assessing the Link Between Premarital Cohabitation and Subsequent Marital Stability." *American Sociological Review*. 53: 127-138.
- [8] Bennentt, Neil G., David E. Bloom, and Cynthia K. Miller (1995) "The Influence of Nonmarital Childbearing on the Formation of First Marriages." *Demography*. 32: 47-62.
- [9] Berkovec, James and Steven Stern (1991). "Job Exit Behavior of Older Men." *Econometrica*. 59(1): 189-210.
- [10] Berry, Steven T. (1992). "Estimation of a Model of Entry in the Airline Industry." *Econometrica*. 60(4): 881-917.

- [11] Breusch, T. and A. Pagan (1979). "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics." *Review of Economic Studies*, 1263-1277.
- [12] Brien, Michael J. (1997) "Racial Differences in Marriage and the Role of Marriage Markets." *Journal of Human Resources*. 32: 741-778.
- [13] Brien, Michael J., Lee A. Lillard, and Steven Stern (1998). "Notes on Cohabitation Model." Unpublished manuscript, <http://www.faculty.virginia.edu/~stevenstern/resint/structstf/blsnotes.pdf>.
- [14] Brien, Michael J., Lee A. Lillard, and Steven Stern (2000). "Technical Appendices for 'Cohabitation, Marriage and Divorce in a Model of Match Quality.'" <http://faculty.virginia.edu/stevenstern/resint/structstf/blsappendix.pdf>.
- [15] Brien, Michael J., Lee A. Lillard, and Linda J. Waite (1999). "Interrelated Family Building Behaviors: Cohabitation, Marriage, and Non-marital Fertility." *Demography*. 36(4): 535-551.
- [16] Browning, Martin, Francois Bourguignon, Pierre Chiappori, and Valerie Lechone (1994). "Income and Outcomes: A Structural Model of Intrahousehold Allocation." *Journal of Political Economy*. 102(6): 1067-1096.
- [17] Bumpass, Larry L., and James A. Sweet (1989). "National Estimates of Cohabitation." *Demography* 26(4): 615-625.
- [18] Bumpass, Larry L., James A. Sweet and Andrew J. Cherlin (1991). "The Role of Cohabitation in Declining Rates of Marriage." *Journal of Marriage and the Family*. 53(4): 913-927.
- [19] Cherlin, Andrew J. (1992). *Marriage, Divorce, and Remarriage*. Cambridge, MA: Harvard University Press.
- [20] Dagsvik, John, Boyan Jovanovic, and Andrea Shepard (1985). "A Foundation for Three Popular Assumptions in Job-matching Models." *Journal of Labor Economics* 3(4): 403-420.
- [21] DeMaris, A. and K. V. Rao (1992). "Premarital Cohabitation and Subsequent Marital Stability in the United States: A Reassessment." *Journal of Marriage and the Family*. 54: 178-190.

- [22] Drewianka, Scott D. (1998). "Commitment and Relationship-Specific Investments: Family Structure." Unpublished manuscript.
- [23] Eckstein, Zvi and Kenneth I. Wolpin (1995). "Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model." *Review of Economic Studies*. 62(2): 263-286.
- [24] Friedberg, Leora (1998). "Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data." *American Economic Review*. 88(3): 608-627.
- [25] Geweke, John (1991). "Efficient Simulation from the Multivariate Normal and Student-t Distributions Subject to Linear Constraints." *Computer Science and Statistics: Proceedings of the Twenty-Third Symposium on the Interface*. 571-578.
- [26] Hajivassiliou, Vassilis (1990). "Smooth Simulation Estimation of Panel Data LDV Models." Unpublished manuscript.
- [27] Harris, Milton and Yoram Weiss (1984). "Job Matching with Finite Horizon and Risk Aversion." *Journal of Political Economy*. 92(4): 758-779.
- [28] Heckman, James (1979). "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica*. 46(6): 931-959.
- [29] Jovanovic, Boyan (1979). "Job Matching and the Theory of Turnover." *Journal of Political Economy*. 87(5, pt. 1): 972-990.
- [30] Keane, Michael (1994). "A Computationally Practical Simulation Estimator for Panel Data." *Econometrica*. 62(1): 95-116.
- [31] Keane, Michael and Kenneth Wolpin (1994). "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." *Review of Economics and Statistics*. 76(4): 648-672.
- [32] Keane, Michael and Kenneth Wolpin (1997). "The Career Decisions of Young Men." *Journal of Political Economy*. 105(3): 473-522.
- [33] Laumann, Edward O., John H. Gagnon, Robert T. Michael, and Stuart Michaels (1994). *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago, IL: The University of Chicago Press.

- [34] Lillard, Lee A., Michael J. Brien, and Linda, J. Waite (1995). "Pre-Marital Cohabitation and Subsequent Marital Dissolution: A Matter of Self-Selection?" *Demography*. 32(3): 437-458.
- [35] Lillard, Lee A. and Linda, J. Waite (1989). "Panel versus Retrospective Data on Marital Histories: Evidence from the PSID." Pp. 243-253 in *Individuals and Families in Transition: Understanding Change Through Longitudinal Data*. U.S. Census Bureau.
- [36] Manski, Charles F. and Steven R. Lerman (1977). "The Estimation of Choice Probabilities from Choice Based Samples" *Econometrica*. 45(8): 1977-1988.
- [37] McFadden, Daniel (1989). "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration." *Econometrica*. 57(5): 995-1026.
- [38] Michael, Robert T. and Nancy B. Tuma (1985) "Entry into Marriage and Parenthood by Young Men and Women: The Influence of Family Background." *Demography*. 22(4): 515-544.
- [39] Milgrom, Paul R. and Robert J. Weber (1982). "A Theory of Auctions and Competetive Bidding." *Econometrica*. 50(5): 1089-1122.
- [40] Miller, Robert A. (1984) "Job Matching and Occupational Choice." *Journal of Political Economy*. 92(6): 1086-1120.
- [41] Moffitt, Robert A., Robert Reville and Anne E. Winkler (1998). "Beyond Single Mother: Cohabitation and Marriage in the AFDC Program." *Demography*. 35(3): 259-278.
- [42] Mortensen, Dale T. (1988) "Matching: Finding a Partner for Life or Otherwise." *American Journal of Sociology*. 94: S215-S240.
- [43] Mroz, Thomas. (1999). "Discrete Factor Approximations in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcomes." *Journal of Econometrics*. 92(2): 233-274.
- [44] Pakes, Ariel. (1986). "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." *Econometrica*. 54(4): 755-784.

- [45] Pissarides, C. A. (1983). "Efficient Job Rejection" *Economic Journal*. 94: 97-108.
- [46] Rindfuss, Ronald R. and A. VandenHeuvel (1990). "Cohabitation: Precursor to Marriage or Alternative to Being Single?" *Population and Development Review*. 16: 703-726.
- [47] Rust, John (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica*. 55(5): 999-1033.
- [48] Rust, John (1988). "Maximum Likelihood Estimation of Discrete Control Processes." *SIAM Journal on Control and Optimization*. 26(5): 1-19.
- [49] Sargent, Thomas J. (1987). *Dynamic Macroeconomic Theory*. Cambridge, MA: Harvard University Press.
- [50] Schmidt, Peter (1982). "Constraints on the Parameters in simultaneous Tobit and Probit Models." In *Structural Analysis of Discrete Data: With Econometric Applications*. edited by Charles Manski and Daniel McFadden. Cambridge, Mass: MIT Press.
- [51] Stern, Steven. (1989). "Estimating a Simultaneous Search Model." *Journal of Labor Economics*. 7(3): 348-369.
- [52] Stinebrickner, Todd R. (1996). "A Dynamic Discrete Choice Model of Teacher Attrition." Ph.D. Dissertation, University of Virginia.
- [53] Teachman, J. D. and K. A. Polonko (1990). "Cohabitation and Marital Stability in the United States." *Social Forces*. 69: 207-220.
- [54] Thornton, Arland, William G. Axinn, and Jay D. Teachman (1995). "The Influence of School Enrollment and Accumulation on Cohabitation and Marriage in Early Adulthood." *American Sociological Review*. 60: 762-774.
- [55] Upchurch, Dawn M., Lee A. Lillard and Constantijn W. A. Panis (1999). "The Impact of Nonmarital Childbearing on Subsequent Marital Formation and Dissolution." Paper presented at the Conference on Nonmarital Fertility, Institute for Research on Poverty, University of Wisconsin - Madison, April 29-30 1999.

- [56] U.S. Department of Health and Human Services (1995). *Report to Congress on Out-of-Wedlock Childbearing*. Hyattsville, Maryland: U.S. Government Printing Office.
- [57] Weiss, Yoram (1997). “The Formation and Dissolution of Families: Why Marry? Who Marries Whom? And What Happens Upon Marriage and Divorce.” In *Handbook of Population and Family Economics*, edited by Mark. R. Rosenzweig and Oded Stark. New York: Elsevier Science.
- [58] Weiss, Yoram and Robert. J. Willis (1985). “Children as a Collective Good and Divorce Settlements.” *Journal of Labor Economics*. 3(3): 268-292.
- [59] Weiss, Yoram and Robert. J. Willis (1997). “Match Quality, New Information, and Marital Dissolution.” *Journal of Labor Economics*. 15(1 Part 2): S293-S329.
- [60] Whittington, Leslie and James Alm (1997). “‘Til Death or Taxes Do us Part.” *Journal of Human Resources*. 32(2): 388-412.
- [61] Willis, Robert. J. and Robert. T. Michael (1994). “Innovation in Family Formation: Evidence on Cohabitation in the United States.” In *The Family, the Market and the State of Aging Societies*, edited by John Ermisch and Naohiro Ogawa. London: Oxford University Press.

D. Tables

Table 1: Logistic Fertility Estimates

Variable	Single	Cohabitation	Married
Black	1.587** (0.079)	0.908** (0.194)	0.081 (0.052)
Catholic	-0.223** (0.095)	-0.090 (0.191)	0.065** (0.034)
South	0.162 (0.103)	0.003 (0.217)	-0.112** (0.041)
Midwest	0.264** (0.105)	0.083 (0.224)	0.068* (0.041)
West	0.133 (0.128)	-0.000 (0.238)	0.098** (0.049)
Years of Education	-0.379** (0.028)	-0.238** (0.048)	0.006 (0.008)
Enrolled in School	-0.686** (0.141)	-0.538 (0.416)	-0.852** (0.096)
Age < 20	0.124** (0.031)	-0.048 (0.150)	-0.296** (0.033)
20 ≤ Age < 23	-0.135** (0.020)	-0.134** (0.048)	-0.023** (0.011)
Age ≥ 23	-0.034** (0.010)	0.026* (0.016)	-0.015** (0.003)
Duration < 1.5 years	0.085** (0.098)	-0.260** (0.106)	0.014 (0.032)
1.5 ≤ Duration < 2.5 years	0.311** (0.085)	0.108 (0.156)	0.038 (0.034)
Duration ≥ 2.5 years	-0.566** (0.123)	-0.018 (0.290)	-0.040 (0.052)

Table 1: Logistic Fertility Estimates (Continued)

Variable	Single	Cohabitation	Married
Number of Children = 0	-1.640** (0.254)	-0.893** (0.423)	-1.871** (0.072)
$1 \leq$ Number of Children < 2	-0.055 (0.174)	-0.101 (0.305)	-0.987** (0.049)
Number of Children ≥ 2	0.106 (0.309)	0.089 (0.609)	-0.300** (0.088)
Youngest Child < 1.5 years	0.648** (0.099)	0.319** (0.161)	0.845** (0.028)
1.5 years \leq Youngest Child < 2.5 years	-0.062 (0.138)	-0.255 (0.344)	0.016 (0.034)
Youngest Child ≥ 2.5 years	-0.262 (0.226)	0.473 (0.594)	-0.517** (0.062)
Constant	-0.659* (0.385)	1.201 (1.335)	0.508* (0.279)

Notes:

1. Standard errors are in parentheses. Double-starred items are significant at the 5% level.
2. Age, duration, and child characteristic effects are piece-wise linear effects.
3. Age, duration, and child age are measured in 6 month time periods.

Table 2: Estimates of Structural Parameters

Variable	Cohabitation	Married
A. Effect on Utility Flows		
Constant	-0.408** (0.007)	0.787** (0.007)
Black	-0.741** (0.010)	-0.673** (0.008)
Catholic	-0.288** (0.007)	-0.266** (0.008)
Years of Education	-0.155** (0.001)	-0.100** (0.001)
Enrolled in School	-2.085** (0.006)	-1.972** (0.006)
South	-0.138** (0.008)	-0.139** (0.006)
Midwest	-0.148** (0.010)	-0.125** (0.008)
West	-0.433** (0.010)	-0.329** (0.008)
Age ≤ 20	0.271** (0.002)	0.171** (0.002)
20 < Age ≤ 23	-0.073** (0.002)	-0.063** (0.002)
Age > 23	0.029** (0.001)	0.014** (0.000)
Duration < 2 years	0.053** (0.003)	0.036** (0.003)
Duration ≥ 2 years	0.101** (0.004)	0.016** (0.003)
Number of Children	0.027** (0.003)	-0.010** (0.002)
Age of Youngest Child	-0.101** (0.002)	-0.139** (0.002)

Table 2: Estimates of Structural Parameters (Continued)

Variable	Cohabitation	Married
B. Effect on Separation Costs		
Constant	2.664** (0.026)	6.728** (0.039)
Number of Children	0.795** (0.014)	1.200** (0.030)
Age of Youngest Child	0.023** (0.005)	0.012** (0.003)
Catholic	1.276** (0.025)	1.861** (0.055)
C. Auxillary Parameters		
σ_θ		0.590** (0.003)
σ_μ		0.065** (0.001)
ρ		0.498** (0.000)

Notes:

1. Standard errors are in parentheses. Double-starred items are significant at the 5% level.
2. Age, duration, and child age are measured in 6 month time periods.
3. Age effects are piece-wise linear effects.
4. Utility flow units are measured relative to the standard deviation of η : $\sigma_\eta = 1.649$.

Table 3: Chi-Square Goodness of Fit Statistics for Simulated Minus Sample Proportions			
Group	DF	Censored χ^2	Normal
Single	514	1443.2	31.61
Cohabiting	160	805.2	39.08
Married	206	220.7	1.02

Notes:

1. Each cell is disaggregated by race (black or white), religion (Catholic or not), education (12, 14, or 16), region (4 regions), and age (16-20,20-24, 24-28, or 28-32).
2. Censored χ^2_1 statistics are censored at the 1% tail (6.63).
3. Normal statistics standardize the censored χ^2 statistics using $N = (\chi^2 - 0.978df) / \sqrt{1.722df}$. The relevant general formula is

$$\begin{aligned}
 E\chi_{1c}^2 &= F_3(c) + c[1 - F_1(c)]; \\
 E(\chi_{1c}^2)^2 &= 3F_5(c) + c^2[1 - F_1(c)]
 \end{aligned}$$

where χ_{1c}^2 is a chi-square random variable with one degree of freedom censored at c and $F_{df}(c)$ is the chi-square distribution function with df degrees of freedom evaluated at c .

Table 4: Heteroskedasticity Tests	
Variance Parameter	Normalized χ^2 -statistic
σ_θ	5.73
σ_μ	-1.20
ρ	1.27
Overall	3.35

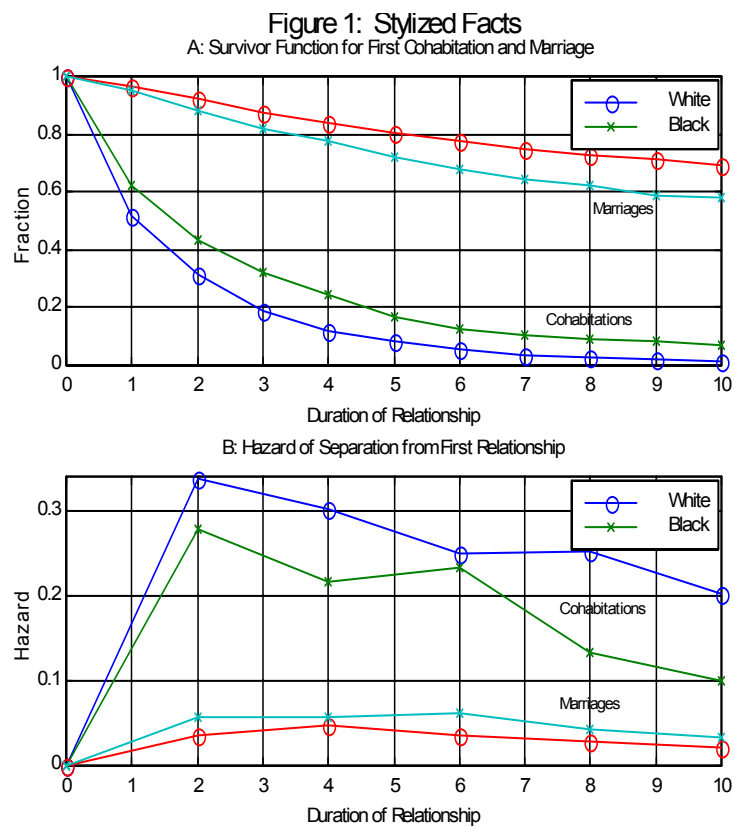
Notes:

1. The χ^2 -statistics have 4 degrees of freedom before normalizing, and, under H_0 : No heteroskedasticity, the normalized statistics are distributed standard normal. The overall statistic has 12 degrees of freedom before normalizing.
2. We test for heteroskedasticity with respect to Black, Catholic, Education, and Region.

Table 5: Analysis of Rare Events

	Single	Cohabiting	Married
A. Single Period Conditional on μ and a Representative Person			
Number of Observations	874,321	40,083	565,069
Frequency of Rare Events	0.07%	10.91%	0.15%
B. Single Period Integrated Over μ and Representative People			
Number of Observations	106,819	6,108	73,991
Frequency of Rare Events	0.005%	0.540%	0.024%
C. Lifetime Paths Integrated Over μ and Representative People			
Number of Observations		6,118	
Frequency of Rare Events		1.16%	

E. Figures



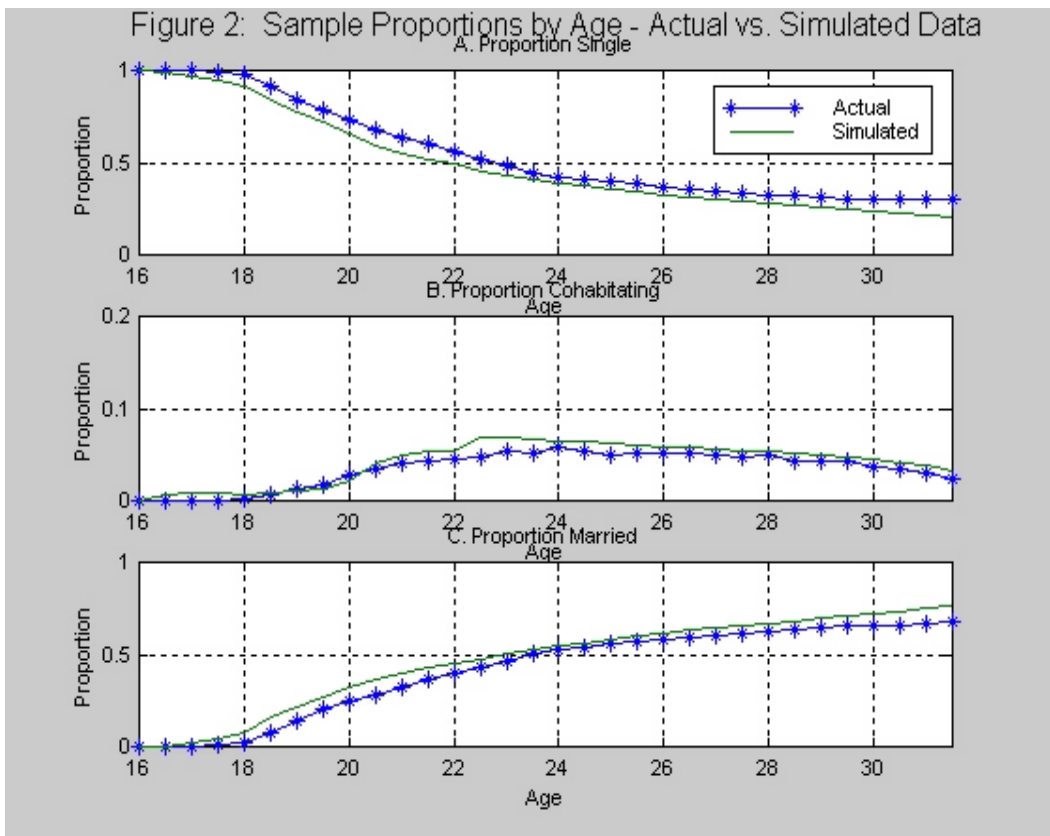


Figure E.1:

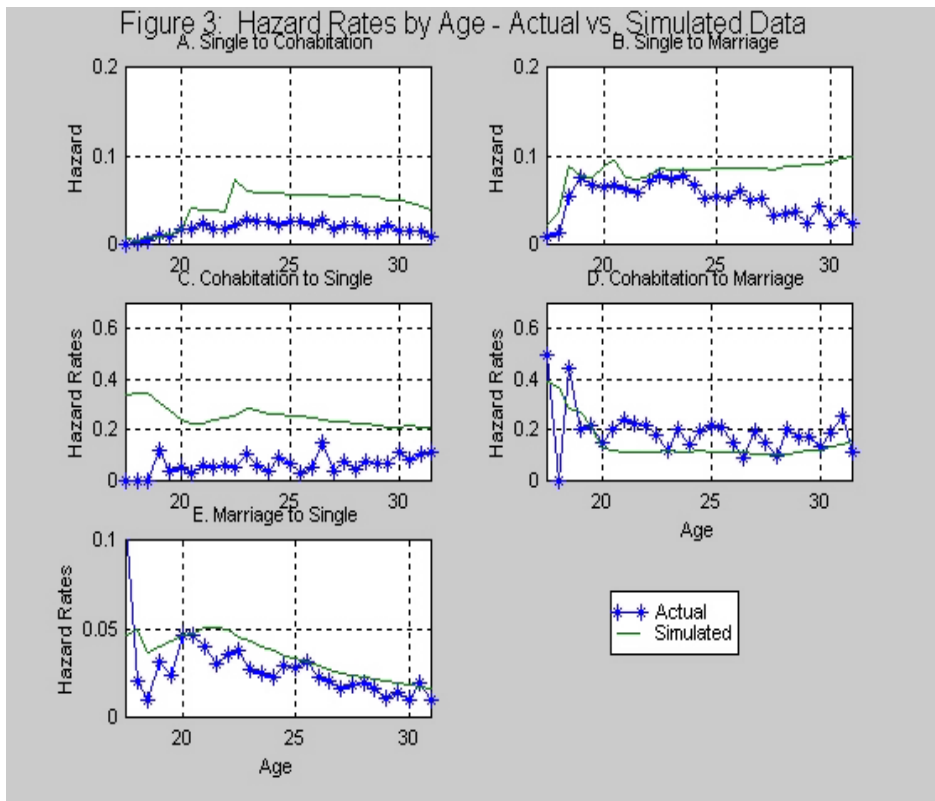


Figure E.2:

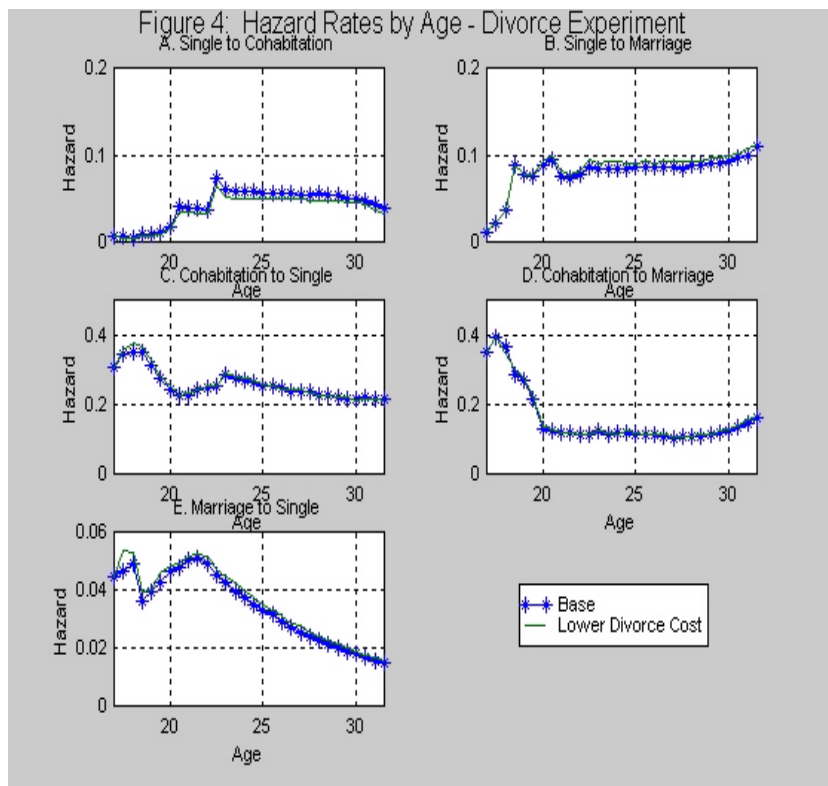


Figure E.3:

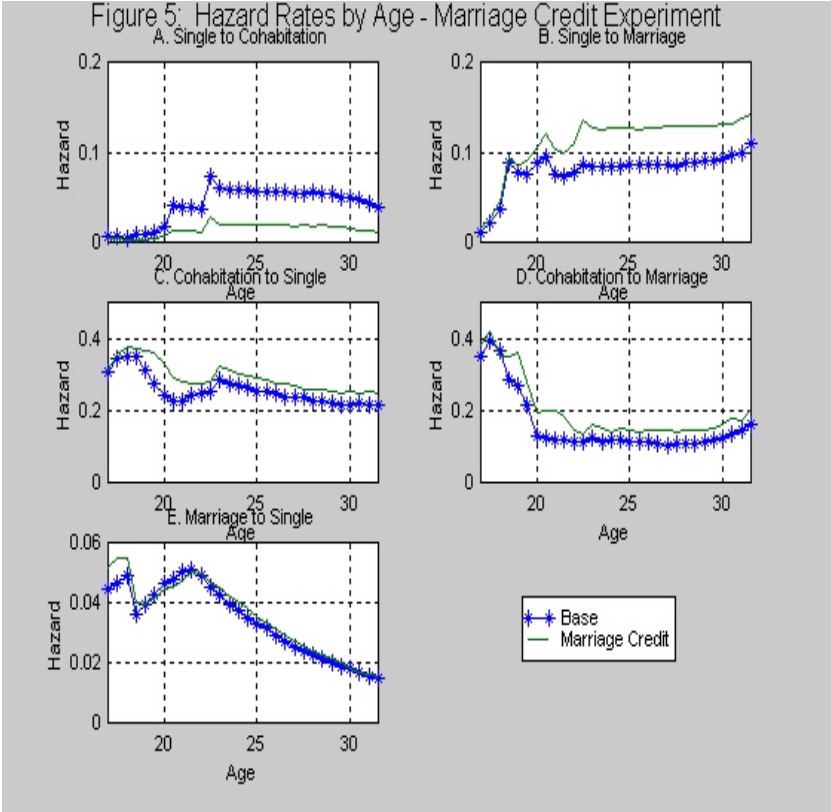


Figure E.4:

Figure 6: Illustrative Weighting Scheme

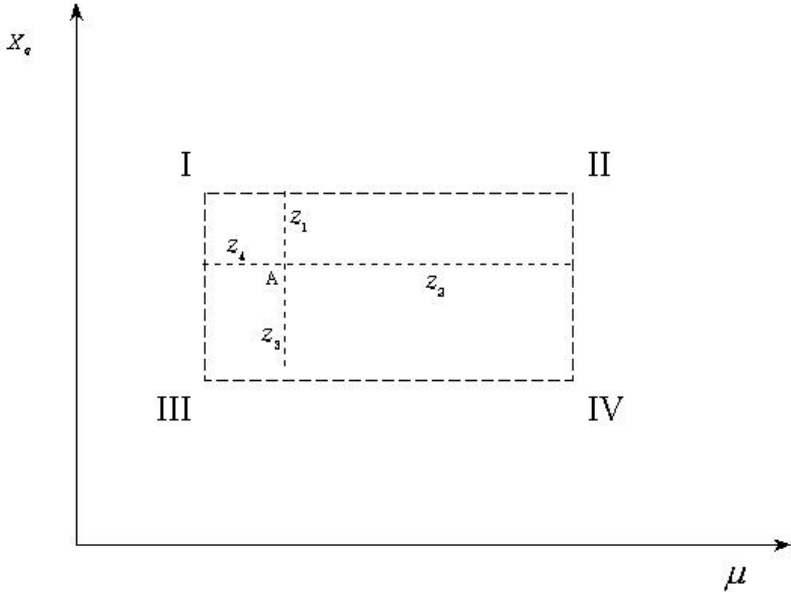


Figure E.5: