Draft: Please Do Not Quote

Projecting Family Formation and Dissolution In a Microsimulation Model

March 2004

Josh O'Harra
Kevin Perese*

Abstract

Dynamic microsimulation models create long-run population projections based on annual micro-level transitions into and out of various demographic and economic states. Within most microsimulation models, a module (or set of modules) exists to create new families and to dissolve existing families. This paper details the set of modules created for the Congressional Budget Office's Long-Term (CBOLT) Model to simulate family formation and dissolution over a 75-year time horizon. The process includes three steps: 1) the selection of individuals, among those at risk of marriage, to become married, 2) given an annual marriage market, the creation of unions with realistic joint-distributions of spousal characteristics via a computationally efficient mate-matching algorithm, and 3) given the characteristics of a married couple and the duration of marriage, the dissolution of marriages through divorce. The estimation techniques, model applications, and results are presented in detail.

*Congressional Budget Office, Washington D.C. 20515. Author contact is Kevin.Perese@cbo.gov. This paper was prepared for the 2004 Annual Meeting of the Population Association of America. We are grateful to members of the CBO Long Term Modeling Group for their contributions to the modeling effort and for specific comments on this paper. The views expressed in this paper are those of the authors and should not be interpreted as those of the Congressional Budget Office.

I. Introduction

Dynamic microsimulation has been used to model phenomena ranging from the spread of AIDS (Morris and Kretzshmar, 2000), to the flow of immigrants (Walker, 1997), to the progression of disease through the human body (Roberts et al., 2000). Though the technique is applicable for numerous (and disparate) exercises, all practitioners seek a common end: to incorporate historical relationships with random possible outcomes to project how the future will unfold. Whether the unit of observation is a civilization, a society, or an individual human being, microsimulation provides one of the of the most promising tools available to for those seeking to model complicated processes over time and into the future.

The contribution offered by this paper is to that tradition of microsimulation concerned with policy analysis. The authors seek to create a longitudinal micro data file that contains information about a synthetic population of future Americans. In particular, we seek information on an individual's age, sex, education, marital status, fertility, health, labor force participation, wealth accumulation, participation in government programs, and, ultimately, their mortality. Although the data set described above will be used to analyze solvency and distributional questions about Social Security, the projections necessary to undertake such analysis rely on demographic analysis. This paper describes one set of building blocks used in the construction of that data set: namely, transitions into and out of marriage and the formation and dissolution of couples.[1]

[1] For expositional simplicity this paper omits a discussion of remarriage, even though it is modeled explicitly. The same techniques described below are used to model this event and the results similar to those presented in Sections III thru V generally obtain.

Microsimulating Marital Transitions for Policy Analysis

Marital transition and couple formation modules exist in several well-established policy simulation micro models. For example, economists at Rand developed the demographic modules for Modeling-Income in the Near Term (MINT), a microsimulation used by the Social Security Administration. For that project, a continuous-time hazard model was used to derive coefficients for a static mircosimulation model that projects Social Security benefits for a cross section of individuals alive in a given base year and through 2040 (Panis and Lillard, 1999).

The Urban Institute's DYNASIM model, last updated in 2000, was constructed to the effect of public policy on issues ranging from welfare, to Social Security, to taxes. Like MINT, it includes marital transitions (modeled as a discrete-time hazard). Unlike MINT, however, DYNASIM also generate spousal linkages. The approach for generating matches in DYNASIM differs somewhat from the one presented here. Rather than relying on econometrically derived correlations, DYNASIM computes an exponential distance function based on differences between age and educational attainment to derive the probability of a union formation.

The last significant microsimulation model considered is the CORSIM (now POLISIM) effort initiated by Steve Caldwell at Cornell University. The marital transition modules in CORSIM rely on a standard logistic regression and utilize variation in age, earnings, labor supply, and education to determine the probability of going into and out of marriage. No explicit attempt is made to capture the pure time effects on marriage and divorce (CORSIM, 2001). With respect to mate matching, CORSIM utilizes a potential pairs data file that consists of a separate observation for every potential match. Each observation contains the characteristics of the potential husband and the characteristics of the potential wife. The mate-matching algorithm in CORSIM begins with an estimation of the relative "compatibility" of each potential couple. The compatibility index is estimated using a logistic regression on a potential pairs data file of recent marriages observed in census data.

The family formation dynamics presented in this paper were developed in the process of constructing the Congressional Budget Office's Long-Term (CBOLT) Policy Simulation Model. Like the models mentioned above, this model was designed to examine the effects of various federal policy alternatives. CBOLT tracks many outcomes by measuring the effect of policy changes on everything from the economy to the federal budget to the Social Security system to individuals. Consequently, CBOLT includes various combinations of actuarial algorithms for projecting population and Social Security system finances, a macro growth model with consumption and labor supply feedbacks, a detailed federal budget accounting framework, and a microsimulation model that operates on a sample of the population. For present purposes, our discussion is limited only to the microsimulation model.[2]

The microsimulation model was first operational in late 2002. In its present form it annually simulates fertility, mortality, immigration, labor force participation, hours worked, earnings, Social Security benefit claiming and, as presented here, marital transitions and pairings. In addition to these processes, current model development efforts are directed toward wealth accumulation (e.g., savings behavior and pensions), micro-level health status and shocks, and micro-level fertility. CBOLT simulates each of

[2] For a complete discussion of the CBOLT model, see O'Harra, et al., (2004)

these events, for each observation, over a projection period that starts in 2003 and ends in 2077.

Like other policy simulation models this model begins with a base dataset; in the case of CBOLT, the starting point is the 1998 Continuous Work History Sample (CWHS).[3] This data set, administered by SSA, contains longitudinal earnings information (from 1951 to 1998) on a 1 percent stratified cluster sample of all people ever issued Social Security numbers, which translates into a sample size of roughly 3 million individuals (Smith, 1989). To convert these data a useable size preparation of the base data file involves an additional 1 in 100 draw, which produces a 1 in 1000 sample of all Social Security Numbers ever issued. The result is that each year, the entire sample includes more than 300,000 individuals recording taxable earnings, total compensation, self-employment status, OASDI benefit entitlement, or death.

The CWHS is preferable for use as a dynamic micro-simulation base-file when compared with the available cross-sectional files such as the Decennial Census or Current Population Survey (CPS) or public-use longitudinal files such as the Panel Survey of Income Dynamics (PSID) or the Survey of Income and Program Participation (SIPP). Available cross-sectional data sets do not have the requisite longitudinal histories needed to project forward using dynamic microsimulation. Publicly available longitudinal data sets like the PSID are much smaller than the CWHS, and those data also suffer from response problems for the highest-earning individuals as well as recall bias. The CWHS is an administrative sample and therefore has more accurate income information for the entire earnings distribution. The downside is that the administrative nature of the CWHS

[3] As of this writing, the base data is being updated to the 2001 CWHS.

limits the demographic data to the information that is available on the initial Social Security number application: year of birth, sex, and race.

Given the lack of detailed demographics, it is necessary to impute several key variables – such as educational attainment, marital status, and marital history – when using the CWHS to project future Social Security taxes and benefits. Those variables are imputed on the basis of measured correlations observed in the PSID, the SIPP, and the CPS. Additionally, part of the imputation strategy relies on a historical simulation using CBOLT's that operates over a period for which there are actual micro data and aggregate control totals with which simulation results can be compared.

The historical simulation exercise is a useful gauge of whether the estimated transition equations are doing a good job of predicting marital outcomes in recent history. The simulations start in 1984, at which point everyone alive in the CWHS has been assigned a marital status, education, and duration in the relevant marital state. The assignments are made such that the starting (1983) population matches the characteristics of the (SSA Area) population that CBOLT uses for its aggregate reference point. At this point, the initial stock of those identified as "married" are matched to spouses. The simulation then proceeds forward from 1984 through 1998 using *actual* CWHS earnings and estimated marital transition equations. The results of the exercise are a set of "calibration factors" which are, in effect, average errors in targeting the observed aggregate marital distributions (again, as reported in the SSA Area Population estimates).

Only after imputations have been made for the historical period and the calibration factors have been derived to correct for systematic biases is the projection of future outcomes possible. The machinery driving these projections is computationally intensive. The model is programmed in Fortran90 to allow for direct operation on multi-dimensional matrices. As of this writing the model consists of nearly 1,000 files and subroutines. The advantage of writing a the model in a "low-level" language is that it allows high-end desktop PCs to evaluate each of the several hundred thousand individuals, all of whom are eligible to experience multiple stochastic demographic and economic transitions in each of 75 years. The result is that solving the dynamic micro model requires 1.4 gigabytes of RAM and, in a world of fixed economic inputs, a solution time of about 20 minutes per run (depending on processor speed).

As mentioned above, the micro model operates in the context of a macro growth framework. By embedding the microsimulation into such an environment, stochastic economic and demographic inputs such as inflation, real wage growth, unemployment, etc. can also be drawn and ranges of probable outcomes (or confidence intervals) can be described.[4] Such a solution is considerably more computationally intensive as it requires solving the micro model multiple times. The result is that 400 runs with randomly drawn inputs require about 150 hours of processing time. These runs are feasible, however, because the simulations can be distributed across multiple machines (for example, machine 1 might run simulations 1 through 50; machine 2 would run 51-100, etc.).

Previous Research on Marital Transitions and Mate Matching

The previous (and ongoing) microsimulation efforts discussed above describe the modeling tradition that CBOLT draws upon and extends. Of course, in constructing a model designed to generate synthetic marital experiences it is necessary to preserve the relationships observed by other social scientists. While this paper introduces novel

[4] For a complete description of the stochastic model used in CBOLT see CBO (2001).

technical applications and extensions its contribution to the social science literature is somewhat limited.

Determinants of marriage and divorce that have been repeatedly identified include age, educational attainment, earnings, duration in married/divorced states, income, the presence or absence of children, and birth cohort (see, e.g. Lillard and Waite, 2000; Peters 1998). These trends are documented and discussed in greater detail in Section II below. Additionally, research on marriage and divorce trends indicate that perhaps purely social perceptions of entering and leaving legally sanctioned unions are changing as well. For example, several scholars have noted the decreasing rate of people ever entering marriage (Raley, 2001; Shoen and Standish 2000; Shoen and Weinick 1993). Demographers have noted that, after increasing for each birth cohort, divorce rates have begun to level off (Goldstein, 1999; Ruggles, 1997).

Understanding mate matching requires knowledge both of the social science related to assortative mating as well as the computer science related to marriage algorithms. In the literature, unions are homogamous along several dimensions; that is people tend to marry people with whom they have a lot in common. For example, education, race, age, socioeconomic status are just a few of the attributes on which couples appear to sort (Qian, 1998; Mare, 1991; Pencavel, 1998).

To operationalize non-random marital sorting and preserve the relationships identified in the economics and sociology literature, scholars have attempted to develop algorithms to match potential pairs. An early effort defined and solved the "stable marriage problem" formally, although an algorithm to match annual pools of residents with hospitals had been in use for more than a decade before the academic publication of the computational algorithm (Gale and Shapley, 1962). A set of "stable marriages" exists if there are no two couples where a partner in *each couple* would prefer to be matched with a person in the *opposite couple*. For this condition to be satisfied, however, what it means to "prefer" one potential mate over another needs to be defined and quantified. Recently, however, the stable marriage algorithm has been criticized for producing too many annual sets of marriages where an exorbitant proportion of the population exhibit husbands that are one year older than the wife (Bouffard et al., 2001). Easther and Vink (2001) suggest that these results are a byproduct of theoretical shortcomings of the stable marriage algorithm. In particular, they argue that the stable marriage algorithm misuses the information contained in the "compatibility" measure estimated for each potential pair.

An alternative to the stable marriage algorithm is to use a stochastic approach to match potential spouses in a marriage market. Rather than relying on an optimization routine to match spouses, a stochastic matching routine processes the information on the basis of Monte Carlo techniques. If the probability of union formation derived in the first step exceeds a random number drawn for a uniform distribution then a marriage occurs. Several policy simulation models incorporate this approach. (Easther and Vink, 2000; Zedlewski, 1990).

II. Data

Microsimulation models often require coefficient estimates based on out-of-sample data. These coefficients serve as the basis for projecting marital transitions. Although this paper explores two phenomena – transitions between marital states and the joint characteristics of spouses – all coefficients are estimated from the marital history topical module in the 1996 SIPP. These data were also linked to longitudinal Social Security earnings records, a proprietary dataset called the Summary Earnings Records (SER)

For simplicity some SIPP categories were collapsed for estimation purposes. For example, the field "married, spouse in home" and "married, spouse out of home" were collapsed into a single variable; similarly, divorced and separated were collapsed into a single "divorced." Even though the sample is a cross-section as of 1996, a longitudinal data file can be constructed because complete marital history profiles can be produced.[5] Although there are possibly 69,000 observations directly available from the SIPP for analysis, the fact that these data are merged with proprietary longitudinal earnings records reduces the sample to 38,380 observations.

These data reveal historical patterns that inform projections into the future. Of particular significance for estimating overall transitions is the evidence that dramatic differences across cohorts are observed, even after controlling for the standard determinants. For example, people born in the 1960s and 1970s are less likely to get married at younger ages (even after controlling for education and other variables) than people born in the 1950s or before.

Although the SIPP is a large and nationally representative sample of the population, there are still reasons for caution when using it to analyze marital transitions, especially over time. Table 1 compares SIPP marital transition rates with aggregate data

---

[5] Because the SIPP tracks data on the most recent marriage, the marriage before the most recent marriage, and the first marriage, there are some marital history profiles where gaps have to be filled in. Because total number of marriages for each individual is known it is possible to fill in "missing marital transitions" for those who have been married 4 times or more. To avoid losing too many observations in the remarriage pool, intervening, unreported marriages were assigned the average duration of all higher order marriages ending in divorce and intervening, unreported divorces were assigned the average duration for a all divorces that end in remarriage.

from the Vital Statistics database collected by the National Center for Health Statistics (NCHS).[6] In general, the aggregate transition rates in the data sets are similar; noticeable differences from the 1990 to 1995 measures include lower divorce rates for both sexes and somewhat lower remarriage rates for men. Those differences can be explained by the slightly different time periods the fact that the SIPP data excludes the institutional population, and reporting problems associated with distinctions between (for example) divorce and separation. The table does not compare SIPP and NCHS marital transitions for earlier time periods, because the SIPP, by construction, only questioned people alive in 1996 about their marital history. Therefore, any comparison of earlier time periods would not be meaningful, but it should be acknowledged that recall bias might affect the analysis that follows.

III. Marriage

As suggested above, understanding the historical trends in marital behavior are changing will be useful in generating projections for the future. Figures 1 through 4 show predicted marital transition probabilities from the SIPP across age and cohort groups. The graphs are produced using a kernel-smoothing technique, where any point on the graph is a weighted average of all observations within a fixed "band" around that age. For example, the value for age 25 in a given cohort is actually a weighted average of people ages 23, 24, 25, 26, and 27 in that cohort, with people age 25 having the highest weights, people ages 24 and 26 having smaller weights, and people ages 23 and 27

[6] The NCHS data is the main data set used by the Social Security Administration in their actuarial analysis of marital transitions. See, for example, Bell (1997). The data were compiled annually from state Vital Statistics offices and issued in numerous public-use formats. However, as a result of budget cuts in 1990, those data are no longer collected or reported

having even smaller weights. All of the vertical axes are expressed as rates per 100 eligible, where eligibility varies with the transition – first marriage eligibles are never married, divorce eligibles are currently married, and remarriage eligibles are currently divorced or widowed.

First marriage rates are shown in Figures 1 through 4. Figures 1 and 2 show annual rates of first marriage, while Figures 3 and 4 show cumulative rates.[7] (One-hundred minus the cumulative rate at any given age is the fraction of people who never marry.) The two sets of figures tell a similar story in slightly different ways. There have been significant declines in rates of first marriage at young ages (through the mid-20s) for all cohorts since the 1950s. However, the *rate* of decline (the gap between the sequentially ordered cohort probabilities at a given age) for both men and women seems to have slowed. Also, there is some evidence that rates of first marriage may actually be slightly higher at older ages (late 20s and older) for more recent cohorts. Thus, as noted, there is no single statement about time/cohort effects that describes what is happening at all ages, which suggests the need for flexibility in the econometric specification for the transition equations.

The 1996 SIPP data used to generate Figures 1 through 4 indicate key aspects about marriage rates that should be considered when estimating marital transition equations: there are generally nonlinear patterns of transition probabilities across age groups *within any given cohort*, and there seem to have been changes in transition rates *at certain ages* across cohorts. It remains to be shown in this section whether those observations can be explained by underlying determinants of marital transition such as education and income (for all transitions) or duration in state (for divorce and

[7] Again, remarriage results are not presented though the process is explicitly modeled.

remarriage). But it is clear that flexibility in the econometric analysis across both age and cohort dimensions is crucial – one would not, for example, want to impose a cohort or time effect that is proportional across all age groups, because some transition probabilities have actually moved in different directions along the age dimension. The SIPP data allow for person-year units of observation. The final set of controls is age, sex, education, lagged income, and time/cohort effects.

Given the list of control variables, the next decision involves choosing an estimation strategy. Most of the exercises above involve splitting the sample by demographics and estimating either logit or hazard models. Although it is clear that one should estimate different equations across sex groups, Figures 1 through 4 indicate that broad age groups will not capture the interesting curvature of the underlying transition probabilities across age groups. Also, it is apparent that cohort/time effects will vary with age, and it is also likely that the effect of some control variables (education or income) will vary across age groups. So, even if one used a polynomial in age to capture the curvature of the transition probabilities by age, the equation would still be imposing the same effect from the other control variables by age.

The econometric approach used here can be thought of as an "age-centered" logistic estimation technique, which is an extension of the group-based approach used by other models. A separate transition equation is estimated for each single year of age and each sex group, but the sample used in the estimation actually includes every observation within a fixed age band around that point. For example, the estimation for males age 25 actually includes all males ages 21 though 29, though, as in the development of the smoothed Figures 1 through 4, the observations farther from the center are weighted less heavily. As a further control, age itself is actually one of the variables used in the estimation. This approach has the desirable aspect that the effect of every other control variable (income, education, duration, and cohort/time effects) is allowed to vary with age (which would be true with separate equations for each age group) but does not suffer from small sample size problems (which would be the case if one ran single equations for each individual age).

The concerns about using an overly restrictive estimation strategy for the transition into first marriage are borne out in the results, which are shown in Tables 2 and 3. Each table shows a series of logits estimated for a given sex group; for example, Table 2 has the results of estimating separate first marriage logits for females ages 17 through 60. [8] Each set of estimated coefficients is reported on a separate row, with significant (at the 90 percent confidence level) variables indicated by **BOLD** numbering. As indicated, there is both a constant and age coefficient for each single year of the age group, which at first blush seems odd but makes sense in the context of the age-centered approach. The age coefficient is an estimate of the slope of the probability function at that age, but it may make more sense to think of the actual simulation "constant" that applies to everyone of a given age as the estimated constant term plus the age coefficient times the value of age.

The predominance of **BOLD** numbering in Tables 2 and 3, especially at the youngest ages, implies that the chosen correlates do in fact significantly affect first

[8] The "bands" actually vary in size at the youngest ages; the equation for 17-year-olds includes 16 through 19, for 18-year-olds includes 16 through 20, etc; for most ages the bands are set to plus or minus four years.

marriage probabilities. For both men and women, education (a dummy variable indicating 14 or more years of shooling) has a significant negative effect on first marriage through the early 20s but then becomes positive and significant for age groups through the mid- to late 30s. Income (which is lagged, not contemporaneous) has a significant positive effect for both women and men at most ages, though the effect turns negative (and eventually becomes significant) for women past their late 40s.

Tables 2 and 3 also suggest that the cohort effects are all significantly negative for the youngest age groups. But, as expected, that effect diminishes with age, as the first marriage probabilities converge for the older age groups (and, as noted, might even be higher for the more recent female cohorts at older ages). The obvious question for forward-looking micro-simulation is how to fill in cohort effects after the historical data ends – that is, for ages 46 and above for the 1950s cohort, ages 36 and above for the 1960s cohort, and ages 26 and above for the 1970s cohort.[9] This paper explores one of several possible approaches.

When considering how to extend the cohort effects, it helps to show the patterns in a graphical form. The solid lines in Figures 5 and 6 show the cohort effects for first marriage (from Tables 2 and 3) by age for the 1950s, 1960s, and 1970s cohorts. The dotted lines show the extensions applied to the cohort terms in order to produce forecasts. Notice first that the last observed cohort term for both males and females (the age 45 value for the 1950s cohort) is assigned to all older age groups for that cohort. This is consistent with the following interpretation: women (men) in the 1950s cohort were

[9] The SIPP data did not suggest any basis for extending trends to the 1980s cohort and beyond, so the last set of residual cohort effects derived in all cases is for the 1970s group. Those residual terms are then applied to the 1980s cohort and all future cohorts "born" into the model.

slightly more (slightly less) likely to get married at age 45 than previous cohorts, and thus it seems reasonable to expect that the same will be true at ages 46 and above.

The second set of observations applies to the values for the 1960s and 1970s cohorts between the ages of 26 and 46. The "s-shaped" pattern of estimated cohort effects is consistent with the shifting of the peak age of first marriage rates to the right over time at the same time the height of the peak is falling (see Figures 1 and 2). By assigning a pattern of cohort terms that is proportionally shifted from the previous cohort, the extensions shown in Figures 5 and 6 keep the shape of the probability distribution constant but continue the shift. Thus, there is a smooth transition of first marriage probabilities at all ages across the three cohorts. The convergence of cohort effects at older ages is the consistent with underlying convergence of the first marriage rates, again, as indicated by the available data in Figures 1 and 2.

IV. Mate Matching

Given that the model generates and annual marriage market, the next step in the simulation of marriage is to unite people in a synthetic household. To estimate the probability of a union forming between any given set of newlyweds, again we use data from the marriage history topical module in the 1996 SIPP. As above, this topical module is linked to longitudinal Social Security earnings records. These same data are also used to provide baseline descriptions of the joint distributions that serve as the benchmark with which simulated outcomes are compared. After dropping observations for missing information, the final analysis file contains 1,277 couples – 834 first marriage couples and 443 higher order marriage couples. [10] Newlywed couples are defined as marriages that started in 1994, 1995, or 1996.

Similar to the characteristics used to analyze the transitions, there are four spousal characteristics that are included in mate-selection models – age, education, earnings, and marriage order. Those variables were selected because of their relevance to retirement and Social Security policy. Social Security rules indicate that age and earnings differentials directly affect the eligibility, benefit calculation, and claiming behavior of retired couples. Although education does not have direct implications for Social Security benefits or eligibility, it is a consistent predictor of mate selection (Mare, 1991; Pencavel, 1998; Qian, 1998). In addition, potential imperfections in the earnings measure make it important to include education in a mate-selection model. Including an education measure in the model allows low-earning, highly educated individuals to be distinguished from low-earning, poorly educated individuals. And finally, because the differences in spousal characteristics are likely to differ by marriage order, these models are estimated separately for first marriages and for higher order marriages.

There are currently no cohort variables included in the mate-matching model. There is some research, however, that the assortative mating patterns with regard to education have changed over time. This phenomenon is largely attributable to the increased educational attainment in the United States and the dependence on schools as marriage markets for nubile singles to find mates (Mare, 1991; Qian, 1998).

[10] Again, while remarriage is specifically modeled in the context of CBOLT the results are excluded here. Technically, the method for forming remarried couples is exactly the same; substantively, the results are somewhat different in that the age, educational, and ALE earnings differentials are all much greater in higher-ordered marriage. This result, based on the SIPP, is maintained by the modeling technique presented in detail below.

Consequently, the joint distribution of spousal characteristics in future unions is assumed to remain the same as those observed in the mid-1990s.

This specification for modeling the likelihood of a match is much richer than the exponential distance function based on age and education used in DYNASIM. This specification, however, is more parsimonious than CORSIM because it aligns on fewer dimensions. Characteristics such as race, fertility, and labor force participation at the time of marriage, which are included in CORSIM, are not included in the models of union formation estimated in this paper because they do not have significant long-run implications on Social Security eligibility, benefit calculation, or claiming behavior.

Age differentials are one of the most salient selection criteria in a mate-matching algorithm. Empirical evidence from the SIPP suggests that there is a nonlinear relationship between spousal age differences and the likelihood of a match. To accurately model this pattern, a combination of age splines and dummy variables are employed in the regression models. For both first marriages and remarriages, dummy variables are used to capture the most likely age differences for marriages. For first marriages, two dummy variables are used – one at an age difference equal to zero, and another at an age difference equal to one. Spline variables in the first marriage model separate the sample by the following age differentials: less than -7 years, -6 to -1 years, +2 to +7 years, and greater than +7 years.

The CBOLT microsimulation model attempts to differentiate between highly educated and poorly educated individuals. Consequently, educational attainment is measured with a dummy variable that equals one if years of education are greater than or equal to 14, and zero otherwise. This limitation may reduce the descriptive power of the mate-matching process, and finer education detail is likely to be included in CBOLT in the future.

To account for potential differences in the joint characteristics of spouses by marriage order, potential pairs data files are created separately for men's first marriages and men's higher order marriages. In addition to estimating the models separately, each model has a dummy variable indicating whether the potential wife's marriage is her first or not.

Finally, a measure of earnings is also included in the model. Empirical evidence from the SIPP indicates that spouses have a tendency to have relatively similar earnings levels. To capture this economic homogamy, individuals are classified into sex-specific quintiles of average lifetime earnings (ALE), and the difference in quintiles is calculated.[11] The difference is specified as husband's ALE quintile minus wife's ALE quintile and ranges from –4 to +4. Both a linear and a squared term are included in the model to capture the quadratic relationship between ALE quintile difference and the likelihood of a match. A quadratic specification assumes that the direction of the difference in the ALE quintiles is not important. That is, marriages in which the man has higher earnings than the wife are just as likely or unlikely as marriages in which the man has lower earnings than the wife.

To estimate the likelihood of a union between potential pairs, we estimate a logistic regression where the probability of union formation is a function of age, education, and earnings differences and marriage order. To estimate this model, a family-level data file that contains both husband and wife characteristics is constructed.

[11] A description of how average lifetime earnings (ALE) are calculated is available from the authors upon request.

For each man in the data file, the characteristics of the current real wife are compared with the characteristics of the *n*-1 other women in the data file of newlyweds. Each comparison is output as a separate record. This process creates a potential pairs analysis file with *n*-squared observations. For every observation where the characteristics of a husband's wife matches those of another newlywed wife, the dependent variable is set equal to one; otherwise, it is set equal to zero. This convention produces a likelihood that bachelor_i would choose bachelorette_j if all he was looking for were the set of characteristics he found in his real wife.

Age, education, average lifetime earnings quintile, and marriage number are the characteristics used to define what a man seeks in a wife. So assume a man entering his first marriage selects a 30-year-old woman who is also in her first marriage, has 16 years of education, and is in the third earnings quintile (relative to other women in the marriage pool). For every observation in the potential pairs data file where this particular man is matched to a woman with those same characteristics (including his actual wife), the dependent variable is set equal to one. This procedure increases the number of marriage events observed in the dependent variable to be greater than *n* in an *n*-squared data file.[12]

Matching Algorithm

The algorithm for mate matching starts with two pools of men and women to be matched in a given year.[13] The first step involves randomly sorting each of the lists of men and women. The next step cuts marriage candidates from their queue if the sizes of the pools are unequal. Excess "marriageables" have their marital status returned to their

[12] Among first marriages, there is a six-fold increase in the number of marriage events observed.
[13] Note that the mate-matching process described here applies only to nonimmigrants. Immigrants in the microsimulation model go through a separate mate-matching process based solely on difference in age.

previous state and are sent back to the general population that will be at risk of marriage again in the following year. This is the only point at which the marital transitions described above are over-written. Removing excess individuals from the marriage queues creates two equally long, randomly sorted lists of men and women to be matched.

The mate-matching process presented here is male centric – each male finds a mate before proceeding to the next male. The low predictive power of the potential pairs model, however, suggests that a matching algorithm would take a very long time and may require several loops over the available females before a match is made. To mitigate this inefficiency, the matching process calculates a normalization factor that is then used to adjust the predicted probabilities.

For each man, a search across all remaining women calculates the normalization factor, which is set equal to the highest predicted probability of a match between him and all the potential women. A second pass through the list of women is when matches are actually determined. For each potential match, a random number is drawn and the predicted probability of the match is divided by the normalization factor. If the adjusted predicted probability is greater than the random number, the match is made. After the match is made, the female is removed from the list of females to be married that year and the algorithm proceeds to the next male to be married. The same steps are repeated until all males have been matched to all females.

The use of the normalization factor ensures that there will be a match made within the second cycle through the available women. The match that has the highest likelihood in the first cycle will have a likelihood of matching equal to one in the second cycle. This produces a similar effect as the methodology implemented in DYNASIM. In

DYNASIM, a male searches over a random selection of 10 available women. If a match is not made on the first pass of those women, then the best match is assigned. In the mate-matching technique employed here, however, the number of women that men search over is *potentially* the entire set of remaining unmarried women. But because the location of the woman that would produce a "match with certainty" is randomly located in the queue, the actual number of women that each man searches over is also random. This technique creates a more randomized process than the one employed in DYNASIM, which arbitrarily limits the search to 10 women for each man before a match is made with certainty.

Results

The matches are based on regression results from men's first and higher order marriage models that are presented separately in Table 4. Almost all of the coefficients are highly statistically significant. Many of the coefficients reveal expected correlations between husband and wife characteristics. Not surprisingly, these results suggest homogamous matchings. Individuals close in age, education, and historical earnings are more likely to marry than individuals who differ by those traits. Similar parameter estimates in columns one and two suggest that the relationship between individual traits and union formation does not differ significantly by marriage order.

Figures 7 through 9 present the results generated by those parameter estimates. In each figure a comparison to the SIPP is included. Figure 7 shows the distribution of simulated spousal age differences against the age differences of newlywed couples observed in the SIPP. Percentage of marriages is plotted on the y-axis and spousal age difference is plotted on the x-axis. The simulated values are based on the annual average distribution of the differences between 2002 and 2076. The distribution plotted for the SIPP is based on marriages that were formed between 1994 and 1996.

One is immediately struck by how well the simulated distribution matches the benchmark distribution in Figure 7. Because the SIPP has considerably fewer marriages than those produced over the 75-year projection period, there is slightly more variation in the distribution of the SIPP characteristics. The figure shows that both the benchmark distribution and the simulated distribution have their peaks at +1. This peak indicates that approximately 15 percent of men's first marriages in the SIPP and over the simulated 75-year projection period are to women that are one year younger. The distribution is fairly even on either side of that peak and approximates a normal distribution.

Figure 8 shows the distribution of education differences produced by the stochastic mate-matching technique discussed above. This distribution indicates that there is strong homogamy by education, which is extensively supported in the research literature (Mare, 1991; Pencavel, 1998; Qian, 1998). The distribution of education differences produced in the simulation closely matches the distribution observed in the SIPP. In each, approximately three-quarters of the couples have the same education level, and regardless of which spouse has more education, the proportions with educational inequities are approximately equal.

Figure 9 shows the simulated distribution of differences in average lifetime earnings quintiles between husbands and wives. The distribution of each appears to be relatively symmetrically distributed, with more than a quarter of the couples marrying spouses that have the same relative economic ranking. Fewer than 5 percent of the marriages in the simulation and in the SIPP have extreme differences between the husbands' and the wives' economic status. This figure shows that the stochastic mate-matching approach produces slightly more marriages characterized by higher earning women (relative to their husbands) than the SIPP might suggest. This pattern may be the result of future compositional changes in marriage markets. As female earnings approach parity with male earnings, there are likely to be ramifications on the characteristics of the matches produced. The overall congruency between the simulated differences in economic status and historically observed differences, however, is satisfactory. Matching along this dimension is particularly important because Social Security spousal benefits and workers' own benefits are affected by relative spousal earnings differentials.

V. Divorce

Unfortunately, not every match lasts forever and, for policy reasons, this matters. To capture divorces we apply several of the same techniques employed in determining marital transitions and extend any observed cohort effects forward. All of the control variables used to model transitions into first marriage are also used to model the dissolution of marriages. In addition, a nonlinear duration function is included in the model.

Figures 10 and 11 show divorce rates by age and cohort for females and males, respectively. For both sexes, divorce rates for cohorts born after 1950 are noticeably higher at younger ages, but those rates are approaching (or even below) the pre-1950 cohort levels in the latest years for which the SIPP data are available. There is also some evidence that divorce rates for some age groups in the 1960s cohort (25 to 35 years old) are actually lower than those of the 1950s cohort, signifying a reversal back toward the pre-1950 cohort rates.

Tables 5 and 6 show age-centered logit estimates for divorce rates of women and men, respectively; again, the importance of flexibility is underscored by the results. Education has mixed (but generally negative) effects on divorce, and income actually has different signs for the two sexes: lower lagged income is more likely to lead to divorce for men, but less likely to lead to divorce for women. The positive or insignificant linear term and negative squared term for duration in many age groups are consistent with underlying divorce patterns: the probability of divorce initially rises after couples are married, but eventually starts to fall after they have been married a certain number of years (though again, the data indicate that this nonlinear effect itself varies with age).

Strong cohort effects are as evident for divorce as they are for first marriage. The cohort-dummy coefficient estimates in Tables 5 and 6 (also shown in Figures 12 and 13) clearly reflect the patterns in the overall divorce probabilities by age and cohort (Figures 10 and 11). Divorce rates are higher for both sexes at younger ages in the 1950s and 1960s cohorts, but that effect disappears with age – there are no significant cohort differences in divorce probabilities for the two younger cohorts at the latest ages which can be observed, and thus one could speculate that divorce rates will match those for the pre-1950 cohorts at older ages.[14]

The dramatic changes in marital transition rates indicated by the SIPP data across age, sex, and cohort groups cannot be explained away by underlying economic variables or changes in educational attainment. Further, the changes are not easily captured by time trends, because the changes in transition rates often go in different directions at different ages, and even the trends that are in a given direction (for example, the drop in

[14] This is consistent with SSA's assumption of an unchanging central (age-adjusted) divorce rate during recent years and for the 75-year forecast horizon.

first marriage rates at young ages) seem to have changed slope, slowing in recent years. To make projections, one can do little more than speculate how those patterns will continue to unfurl for existing cohorts as they age and for future cohorts at all ages. The implications of the choices made for extending the cohort terms are drawn out in the next section, which uses the estimated transition equations in a dynamic micro-simulation setting.

VI. Microsimulation Results

The goal of this paper is to show how to generate a set of longitudinal marriage histories for a *future* sample of the population in a dynamic micro-simulation context. The sections above lay out the modeling process for projecting marital transitions in three distinct steps: entering into an annual marriage market, forming new couples through a mate-matching model, and dissolving unions. This section describes how those pieces are brought together in the Congressional Budget Office Long Term (CBOLT) policy simulation model and provides some basic results of the marital projection over the 2000 to 2075 period.[10] The results depend on a variety of assumptions, including how transition probabilities will evolve in the future. The illustrative calculations presented here show the implication of just one set of assumptions.

As mentioned above we employ our projection equations to estimate transitions for a period where the actual outcomes are known; in this case, we model the events between 1984 and 1998. This approach allows us to determine if the transition equations have been properly programmed into the model and to ensure that the properties of our base data file (the CWHS) are consistent with the properties of the estimation data file (the SIPP). Because slight systematic errors result by age, sex and transition type, we derive a set of calibration factors from the historical simulation period (1984-1998) to use throughout the projection (2001-2076). These calibrations are presented in Figures 14 and 15. The figures reveal that for most age and sex groups the adjustments to underlying transition probabilities vary by fewer than 5 percentage points for women and closer to 2 percentage points for men. Because there are so few transition events at older ages, the estimates perform relatively poorly for those above 70 and the resulting calibration factors are somewhat larger.[15]

The fact that the marital transition equations estimated here are able to capture the trends in aggregate marital distributions during the 1984 to 1998 period suggests that (setting aside uncertainty about how to extend the cohort terms) the model is capable of generating reasonable cross-sectional marital distributions. That conclusion is borne out in Table 7 for the entire population, and in Figures 16 and 17 for people ages 62 to 67 for whom Social Security projections are particularly relevant. At this point, we seek to ensure that our projections are reasonable. To that end, we compare our cross-sectional micro-based results with actuarial projections from the Office of the Chief Actuary at the Social Security Administration (OCACT).

Table 7 shows cross-sectional marital distributions over time in both the OCACT and CBOLT base case projections. The most striking result is the similarity between the cross-section marital distributions over the 75-year period; but closer inspection shows a few small but systematic differences.[16] For example, our approach consistently produces

[15] Recall that marital transitions at these ages are also based on the AHEAD dataset which may differ more from the CWHS than does the SIPP.

[16] The similarities are even more striking when one considers that some of the processes used to project forward are very different. For example, OCACT uses the distribution of "available" mates by age and sex to predict how many marriages will occur, while CBOLT does not. For a thorough discussion of the OCACT actuarial projections, see Bell (1997).

fewer divorcees as a proportion of the population. Also, by the end of the projection period, our estimates yield about the same percentage of married females but a slightly higher fraction of never married males. The proportion widowed is slightly different, but that result stems from different mortality projections.

At first glance the fact that the proportion of the population that never married is decreasing while the proportion married is increasing might appear to contradict the earlier story suggesting that fewer people are marrying. However, because the age composition of the population is changing, measures of overall cross-sectional distribution by marital status do not accurately reflect the trends in either the OCACT or CBOLT projections. Figures 16 and 17 isolate the trends in the cross-sectional marriage distribution for people ages 62 through 67. That age group is interesting because, by this point in their lives, most of the cumulative effects of all three transitions modeled in this paper will probably be evident. These graphs suggest that the age-centered transition equations, combined with the minor adjustment factors derived from the historical simulations, generate individual behaviors that (when added up) approximately match the aggregate projections used by the Social Security actuaries. Furthermore, these pictures support our hypothesis that the transition equations should generate a smooth transition to a new distribution of marital status outcomes.

The CBOLT marital transition equations seem capable of generating reasonable cross-sectional patterns, but what about longitudinal outcomes? In addition to classifying the right number of people as marrying, divorcing and remarrying the right number of people in any given year, it is important to ensure that these transitions occur at the appropriate points in an individual's life-cycle and that the distributions of the number of marriages and marital durations across cohorts are reasonable.

The first bit of evidence that the longitudinal properties are satisfactory comes from looking at mean age at first marriage, shown in Figure 18. The solid lines represent Vital Statistics data covering the period between 1964 and 1990. The CBOLT projections start in 1984 and extend through 2075, thus there is a six-year overlap in the graphs. Historically, the mean age at first marriage has been drifting upward, which is consistent with the cohort analysis presented in Figures 1 and 2. Between 1984 and 1990, the CBOLT historical simulation produced mean ages of first marriage very near those registered in the Vital Statistics. That similarity suggests that the model accurately predicts the timing of transitions into first marriage. While the trend of delaying first marriage is predicted to continue for some time, it eventually stabilizes at about age 28 for both sexes, because once the 1980s cohort is through the relevant ages there are (by construction) no further cohort effects. Note that the average age differential between the sexes shrinks over time, which is consistent with the extended positive cohort term for female first marriage probabilities (and negative term for males).

The second observation on longitudinal outcomes comes from looking at the trend in marital duration at the time of divorce, as shown in Table 8. This statistic is particularly interesting to Social Security analysts because of the duration requirement that governs benefit eligibility for divorced spouses. The CBOLT transition equations suggest that the typical duration of those marriages ending is divorces is not necessarily expected to fall. Again, at first glance, this result may seem counterintuitive, because of increased divorce rates. This trend is, however, supported both by actual data and theoretical determinants of divorce. Vital Statistics reports a steady, albeit slight, increase in the duration of marriages ending in divorce. This increase suggests a smaller proportion of very short marriages. Divorce researchers have posited that as marriage ages drift up and cohabitation rates increase, the number of weak, official marriages will decline (Bumpass and Sweet, 1989).

VII. Policy Implication and Conclusions

There are ultimately many ways to tabulate longitudinal marriage outcomes in a dynamic micro-simulation context and thus draw conclusions about the model's properties. When trying to understand a model it sometimes helps to ask a policy question and see what the model offers. For example, one interesting question about marriage in the Social Security context is, What will happen to the number of women eligible for spousal benefits? That question is policy relevant insofar as elderly women have historically been vulnerable to high poverty rates, but recent trends in female labor force participation are poised to ameliorate the condition of future cohorts (Social Security Administration, 2000).[17] The longitudinal marital techniques presented in this paper and implemented within the context of a larger micro simulation permit an illustration of how the distribution of marital outcomes varies across women with different lifetime income.

Table 9 addresses the above question. It presents the percentage of *nonwidowed* women ages 62 to 67 who are eligible to draw OAI spouse benefits. The relevance of controlling for widowhood warrants a brief explanation. As male mortality rates decline,

[17] Labor force participation is also forecast; currently it is modeled as a nested logit where participation is a function of age, sex, marital status, birth cohort, and beneficiary status.

the number of husbands increases, thus increasing overall spouse eligibility. This "mortality improvement" confounds the effect of the retreat from marriage described above. Once the marital status composition is controlled for, two trends stand out. The first, going down a given column, suggests that women with lower lifetime earnings are more likely to be eligible for spouse benefits than are their richer counterparts. In 2005, 95% of women in the lowest lifetime earnings decile will be covered by their spouse's earnings while only 77% of women in the highest decile will be comparably covered. This paper does not address whether these women stay married because they are likely low earners or if they are low earners because they know they are likely to stay married. The second trend, reading the rows across, suggests that the probability of a low-earning woman losing OAI spouse coverage is greater than the probability of a high-earning woman will losing that coverage. The estimated change in coverage among the lowest decile of nonwidowed women is eight percentage points, from 95 percent in 2005 to 87 percent by 2075. The comparable decline for women in the highest decile is only five percentage points.

References

Bell, Felicitie C. *Social Security Area Population Projections: 1997.* Actuarial Study #112, Social Security Administration, Office of the Chief Actuary. SSA Pub. No. 11-11553. (August 1997).

Bouffard, Neal, Richard Easther, Tom Johnson, Richard J. Morrison, and Jan Vink. 2001. “Matchmaker, Matchmaker, Make Me a Match.” *Brazilian Electronic Journal of Economics*. Vol. 4, No. 2.

Bumpass, Larry L. and John A. Sweet. “National Estimates of Cohabitation,” *Demography*, vol. 26, no.4, pp. 615-625 (November 1989).

Caldwell, Steven, Melissa Favreault, Alla Gantman, Jagadessh Gokhale, Thomas Johnson, and Laurence J. Kotlikoff. “Social Security’s Treatment of Postwar Americans.” NBER Working Paper #6603 (June 1998).

Congressional Budget Office. *Uncertainty in Social Security’s Long-Term Finances: A Stochastic Analysis*. CBO Paper (December 2001).

Easther, Richard and Jan Vink. 2000. “A Stochastic Marriage Market for CORSIM.” Strategic Forecasting Technical Paper: http://www.strategicforecasting.com/cgi-bin/filedesc.cgi?filename=REJVmarriage_101000.ps&path=REasther_paper/

Frees, Edward W. (Jed). “Summary of Social Security Administration Projections of the OASDI System,” Working Paper for the 1999 Technical Panel on Assumptions and Methods, Social Security Advisory Board (December, 1999).

Gale, D. and L.S. Shapely. 1962. “College Admissions and the Stability of Marriage.” *American Mathematical Monthly*, 69:9-15.

Goldstein, Joshua R. “The Leveling of Divorce in the United States,” *Demography*, vol. 36, no. 3, pp. 409-414 (August 1999).

Gusfield, Dan and Robert W. Irving. 1989. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press: Cambridge, Mass.

Harris, Amy Rehder, and John Sabelhaus. “Projecting Longitudinal Earnings for Long-Run Policy Analysis,” Congressional Budget Office Technical Paper, 2003-02 Washington, D.C. 2003.

Johnson, Tom. 2000. “Stable Marriages vs. Optimal Marriages.” Strategic Forecasting Technical Paper: http://www.strategicforecasting.com/cgi-bin/filedesc.cgi?filename=TJstable_102400.ps&path=TJohnson_paper/

Knuth, Donald E. 1976. *Marriages Stables*. Les Presses de l'Université de Montréal: Montréal, Quebec, Canada.

Kreider, Rose M. and Jason M. Fields. *Number, Timing, and Duration of Marriages and Divorces: Fall 1996*. Current Population Reports, P70-80. U.S. Census Bureau, Washington, D.C. (2001).

Lillard, Lee, and Stan Panis. "Demographic Projections." Paper presented at First Annual Joint Conference for the Retirement Research Consortium, "New Developments in Retirement Research" (May 20-21 1999).

Lillard, Lee and Linda J. Waite. "Marriage and the Work and Earnings Careers of Spouses." Paper presented at the Second Annual Joint Conference for the Retirement Research Consortium, "The Outlook for Retirement Income" (May 17-18, 2000).

Mare, Robert D. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Review*, 56:15-32.

Morris, Martina and Mirjam Kretzschmar. "A Microsimulation Study of the Effect of Concurrent Partnerships on the Spread of HIV in Uganda." Population Research Institute, Working Paper 00-07. University Park, PA (June, 2000).

O'Harra, Josh, John Sabelhaus, and Michael Simspon. "Overview of the Congressional Budget Office Long-Term (CBOLT) Policy Simulation Model," Congressional Budget Office, Technical Paper 2004-1. Washington, D.C. (2004).

Orcutt, Guy H., Steven Caldwell, and Richard Wertheimer II. 1976. *Policy Exploration Through Microanalytic Simulation*. Washington, DC: Urban Institute Press.

Panis, Constantijn, and Lee Lillard. "Final Report: Near Term Model Development, Part II." Rand Corporation, Santa Monica, Calif. (August 1999).

Pencavel, John. 1998. "Assortative Mating by Schooling and the Work Behavior of Wives and Husbands." *American Economic Review*, 88:326-329.

Perese, Kevin. "Documentation for the Divorce Module." The Urban Institute, Washington, D.C. (1999).

Peters, H. Elizabeth. "Retrospective Versus Panel Data in Analyzing Lifecycle Events," *The Journal of Human Resources*, vol. 23, no. 4, pp. 488-513 (Fall 1988).

Preston, Samuel H., and John McDonald. "The Incidence of Divorce Within Cohorts of American Marriages Contracted Since the Civil War." *Demography*, vol. 16, no. 1, pp. 1-25 (February 1979).

Raley, R. Kelly. “Recent Trends and Differentials in Marriage and Cohabitation: The United States.” *The Ties that Bind*, ed., Linda Waite. New York: Aldine de Gruyter (2000): 19-39.

Qian, Zhenchao. 1998. “Changes in Assortative Mating: The Impact of Age and Education, 1970 - 1990.” *Demography*, 35:279-292.

Ruggles, Steven. “The Rise of Divorce and Separation in the United States, 1880-1990.” *Demography*, vol. 34, no. 4, pp. 455-466 (November 1997).

Roberts, M., C.L. Bryce, and D.C. Angus. “Predict Natural History by Survival or Disease Progression but not Both: Inconsistencies in Monte Carlo Microsimulation.”

Schoen, Robert, and Robin M. Weinick. “The Slowing Metabolism of Marriage: Figures from 1988 U.S. Marital Status Life Tables.” *Demography*, vol. 30, no. 4, pp. 737-746 (November 1993).

Schoen, Robert, and Nicola Standish. *Footprints of Cohabitation: Results from Marital Status Life Tables for the U.S., 1995.* Working Paper 00-12, Population Research Institute, University Park, Penn (September 2000).

Smith, Creston M. “The Social Security Administration’s Continuous Work History Sample.” *Social Security Bulletin* 52, 10 (October 1989): 20-28.

Social Security Administration. *Income of the Population 55 or Older, 2000*. Washington, D.C.

Walker, Agnes. “Modelling Immigrants to Australia – to Enter a Dynamic Microsimulation Model.” Presented at the *International Conference on Combinatorics, Information Theory, and Statistics*. Portland, MA. (July 18-20, 1997).

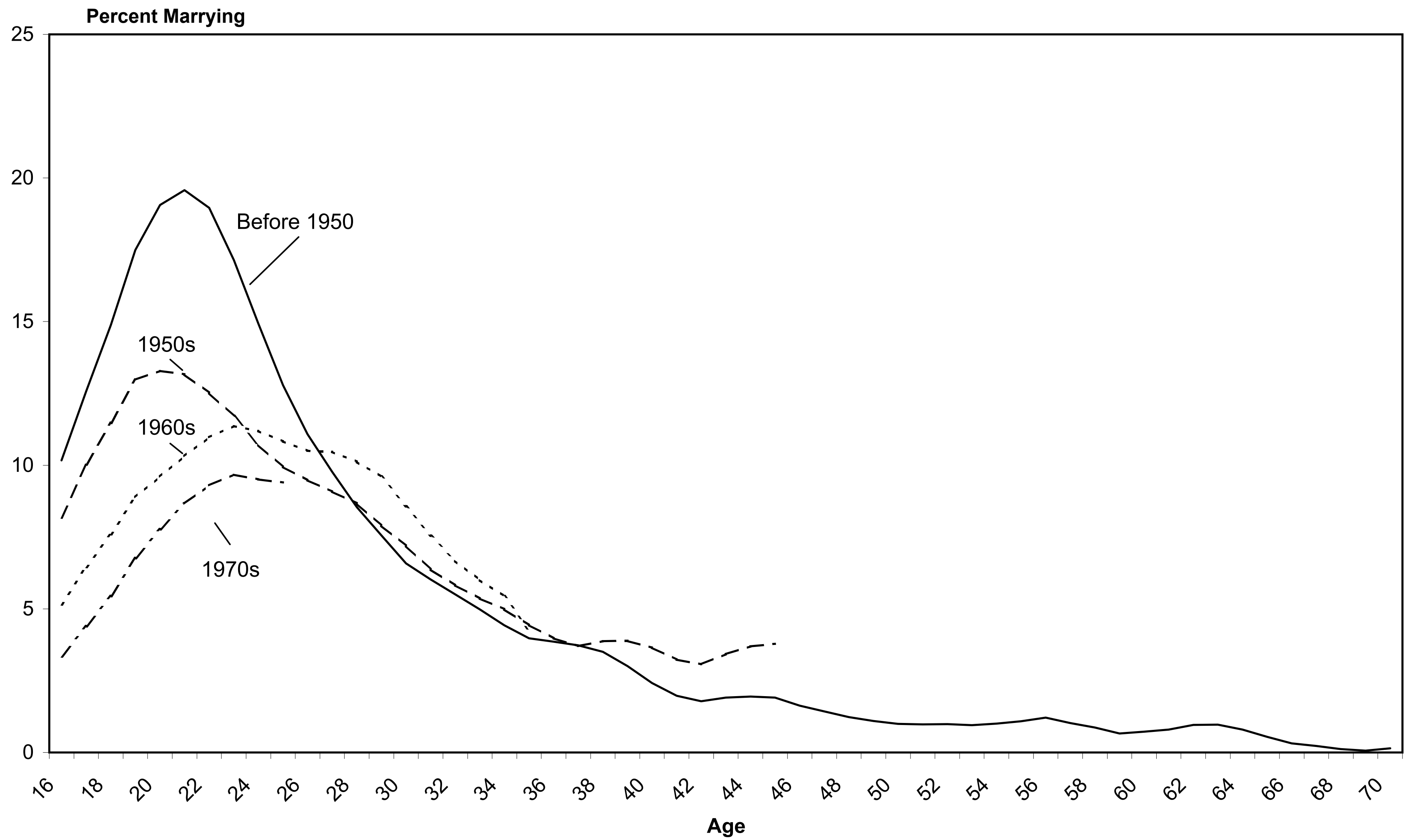
Zedlewski, Sheila R. 1990. “The Development of the Dynamic Simulation of Income Model (DYNASIM).” In Gordon H. Lewis and Richard C. Michel, Editors. *Microsimulation Techniques for Tax and Transfer Analysis*. Washington, DC: Urban Institute Press.

Table 1.
Comparison of Marital Transition Rates in SIPP and Vital Statistics (NCHS) Data

	Men	Women
First Marriage Rate (Per 1,000 People)		
SIPP (1995)	48.6	58.3
NCHS (1990)	47.0	57.7
Divorce Rate (Per 1,000 People)		
SIPP (1995)	14.5	17.4
NCHS (1990)	19.2	18.7
Remarriage Rate (Per 1,000 People)		
SIPP (1995)	73.2	38.9
NCHS (1990)	84.5	35.8

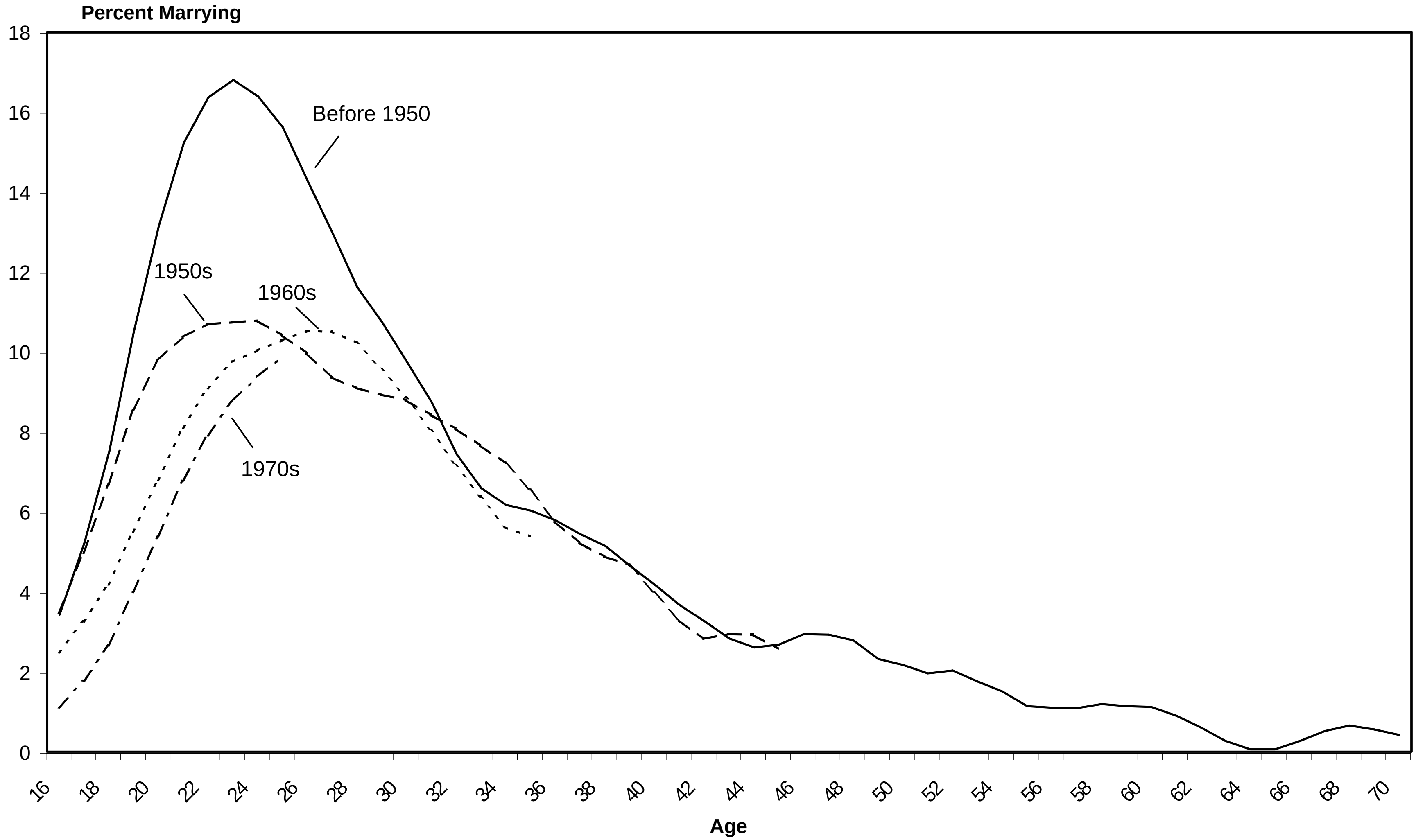
Note: SIPP = Survey of Income and Program Participation
NCHS = National Center for Health Statistics

Figure 1: Female First Marriage Rates by Cohort and Age



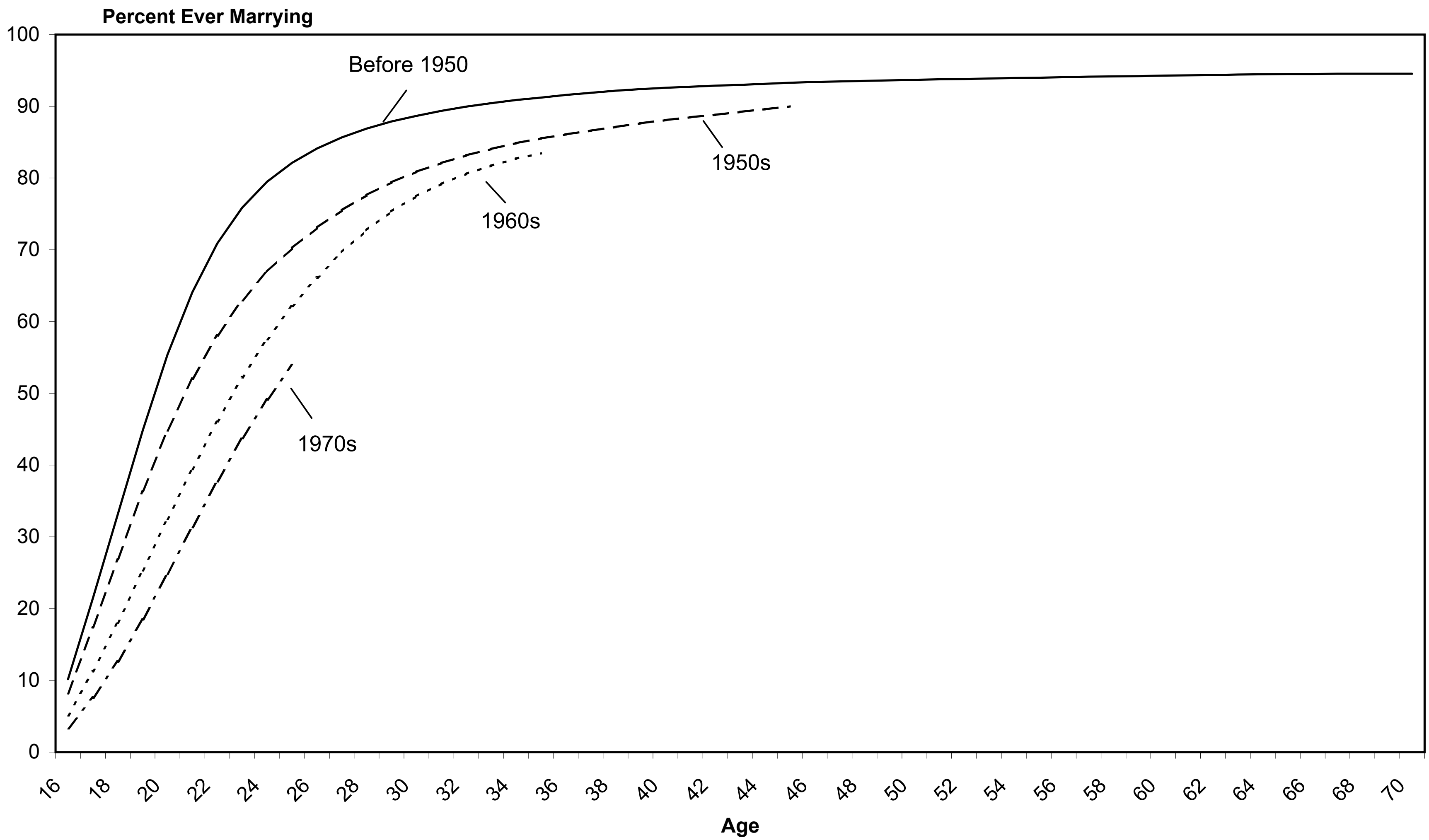
Source: Congressional Budget Office based on data from the 1996 Survey of Income and Program Participation.

Figure 2: Male First Marriage Rates by Cohort and Age



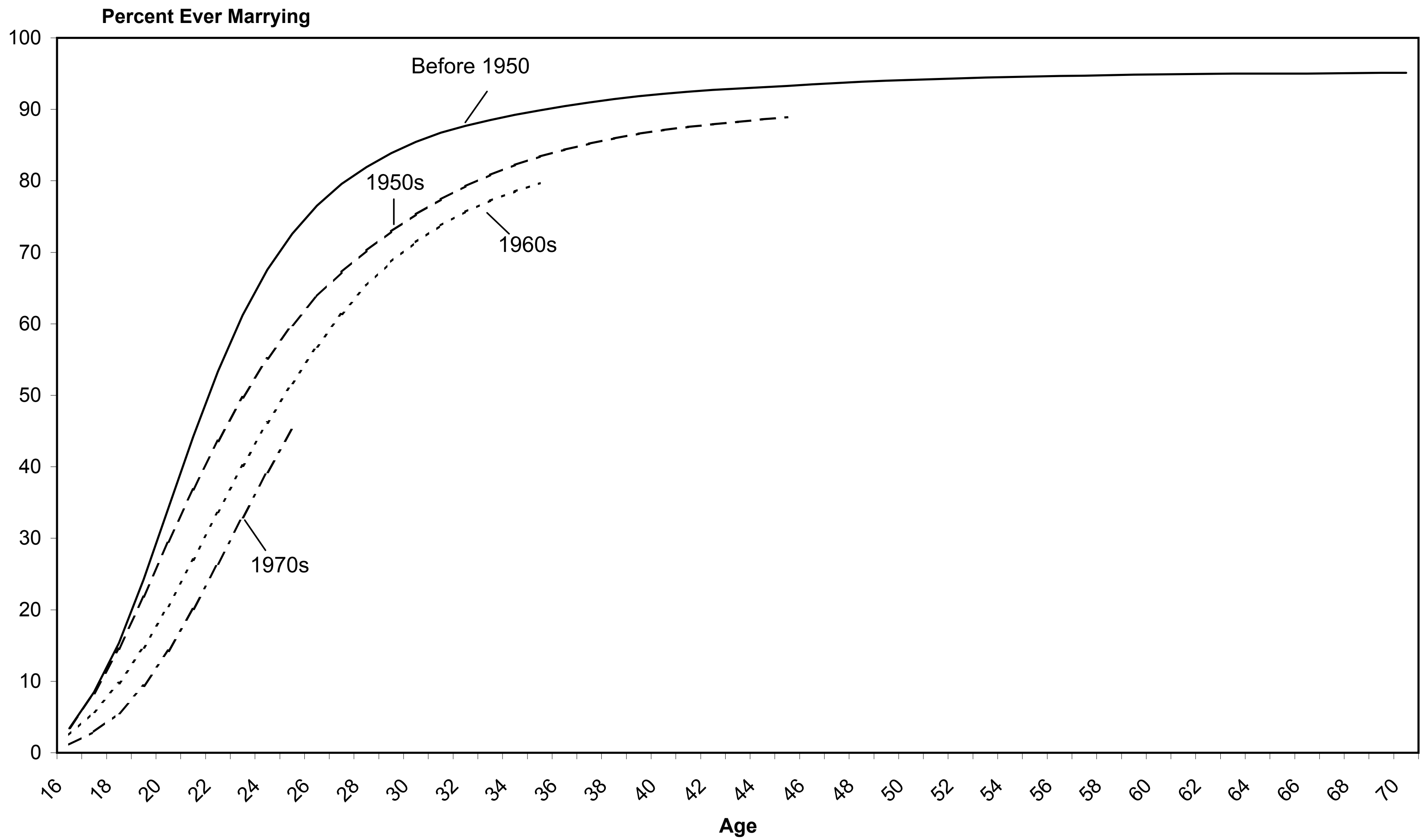
Source: Congressional Budget Office based on data from the 1996 Survey of Income and Program Participation.

Figure 3: Female Cumulative First Marriage Probabilities by Cohort and Age



Source: Congressional Budget Office based on data from the 1996 Survey of Income and Program Participation.

Figure 4: Male Cumulative First Marriage Probabilities by Cohort and Age



Source: Congressional Budget Office based on data from the 1996 Survey of Income and Program Participation.

Table 2
Age-Centered Regression Coefficients, Female First Marriage

Age Group	**Coefficient on:**						
	Age	**Educ**	**Income**	**Cohort50**	**Cohort60**	**Cohort70**	**Constant**
17	**0.4078**	**-1.2836**	**0.0505**	**-0.1849**	**-0.7488**	**-1.3069**	**-8.8695**
18	**0.1984**	**-0.9851**	**0.0646**	**-0.2654**	**-0.7750**	**-1.2554**	**-5.3802**
19	**0.1180**	**-0.6605**	**0.0687**	**-0.3493**	**-0.7948**	**-1.1919**	**-4.0094**
20	**0.0659**	**-0.4325**	**0.0692**	**-0.4082**	**-0.7761**	**-1.1156**	**-3.1019**
21	**0.0132**	**-0.2548**	**0.0652**	**-0.4492**	**-0.7397**	**-1.0336**	**-2.0682**
22	**-0.0232**	**-0.0900**	**0.0618**	**-0.4828**	**-0.6770**	**-0.9533**	**-1.3510**
23	**-0.0532**	**0.0513**	**0.0564**	**-0.4876**	**-0.5919**	**-0.8607**	**-0.7396**
24	**-0.0727**	**0.1578**	**0.0496**	**-0.4645**	**-0.4916**	**-0.7593**	**-0.3318**
25	**-0.0838**	**0.2281**	**0.0432**	**-0.4024**	**-0.3601**	**-0.6494**	-0.1167
26	**-0.0840**	**0.2624**	**0.0390**	**-0.3151**	**-0.2245**		-0.1754
27	**-0.0793**	**0.2945**	**0.0362**	**-0.2254**	**-0.1026**		**-0.3651**
28	**-0.0835**	**0.3008**	**0.0346**	**-0.1457**	-0.0056		-0.3108
29	**-0.0940**	**0.3058**	**0.0337**	**-0.0877**	0.0742		-0.0644
30	**-0.1042**	**0.2852**	**0.0340**	-0.0342	**0.1274**		0.2044
31	**-0.1142**	**0.2674**	**0.0309**	-0.0095	**0.1480**		0.5204
32	**-0.1114**	**0.2229**	**0.0295**	0.0093	**0.1486**		0.4556
33	**-0.1080**	**0.1985**	**0.0257**	0.0166	0.1408		0.3846
34	**-0.1009**	**0.1816**	**0.0224**	0.0332	0.1040		0.1813
35	**-0.0910**	**0.1917**	**0.0157**	0.0290	0.0497		-0.1097
36	**-0.0733**	0.1938	**0.0135**	0.0284			-0.7145
37	**-0.0724**	0.2457	**0.0068**	0.0432			-0.7342
38	**-0.0822**	0.2901	**0.0034**	0.0897			-0.3866
39	**-0.0928**	0.2675	**0.0022**	0.1479			0.0042
40	**-0.1096**	0.2058	0.0061	**0.2174**			0.6326
41	**-0.0974**	0.1301	0.0085	**0.3348**			0.1209
42	**-0.0704**	-0.0180	0.0147	**0.4543**			-1.0136
43	-0.0275	-0.2326	0.0202	**0.5487**			**-2.8201**
44	-0.0171	**-0.4013**	0.0197	**0.6266**			**-3.2628**
45	-0.0426	**-0.5414**	0.0165	**0.6478**			-2.0977
46	**-0.1045**	**-0.7567**	-0.0020				0.8374
47	**-0.1456**	**-0.7762**	-0.0145				2.8212
48	**-0.1245**	**-0.6992**	-0.0240				1.8959
49	**-0.1060**	-0.4757	-0.0248				0.9858
50	-0.0741	-0.4092	-0.0412				-0.5410
51	-0.0463	-0.3358	-0.0530				-1.8714
52	0.0067	-0.2539	-0.0525				-4.5669
53	0.0246	-0.2141	-0.0578				-5.5385
54	0.0383	-0.2991	**-0.0756**				-6.1978
55	-0.0063	-0.2090	**-0.0763**				-3.7633
56	-0.0514	-0.0590	**-0.0762**				-1.2904
57	-0.0860	-0.1108	**-0.0827**				0.6373
58	-0.1177	-0.2401	**-0.0986**				2.5872
59	-0.1025	-0.3964	**-0.0975**				1.7546
60	-0.0212	-0.6622	**-0.0912**				-3.0412

Note: Coefficients in **BOLD** are significant at the 90% level.

Table 3
Age-Centered Regression Coefficients, Male First Marriage

Age Group	Coefficient on: Age	Educ	Income	Cohort50	Cohort60	Cohort70	Constant
17	**0.5626**	**-1.0578**	**0.0778**	0.0000	**-0.4407**	**-1.2186**	**-12.8694**
18	**0.3540**	**-0.8410**	**0.0891**	**-0.1243**	**-0.6510**	**-1.1558**	**-9.2769**
19	**0.2512**	**-0.5673**	**0.0890**	**-0.2672**	**-0.7204**	**-1.0829**	**-7.4000**
20	**0.1761**	**-0.3767**	**0.0859**	**-0.3591**	**-0.7265**	**-1.0163**	**-5.9724**
21	**0.1187**	**-0.2460**	**0.0793**	**-0.4191**	**-0.7196**	**-0.9445**	**-4.7703**
22	**0.0705**	**-0.1339**	**0.0740**	**-0.4628**	**-0.6910**	**-0.8824**	**-3.7333**
23	**0.0277**	**-0.0372**	0.0680	**-0.4944**	**-0.6512**	**-0.8263**	**-2.7730**
24	-0.0062	**0.0453**	**0.0626**	**-0.4967**	**-0.5944**	**-0.7504**	**-1.9937**
25	**-0.0316**	**0.1084**	**0.0585**	**-0.4842**	**-0.5204**	**-0.6805**	**-1.3986**
26	**-0.0497**	**0.1596**	**0.0550**	**-0.4417**	**-0.4370**		**-0.9735**
27	**-0.0590**	**0.2023**	**0.0534**	**-0.3964**	**-0.3607**		**-0.7759**
28	**-0.0639**	**0.2371**	**0.0537**	**-0.3281**	**-0.2780**		**-0.7048**
29	**-0.0717**	**0.2678**	**0.0535**	**-0.2575**	**-0.2134**		**-0.5385**
30	**-0.0765**	**0.2896**	**0.0522**	**-0.1731**	**-0.1607**		**-0.4452**
31	**-0.0782**	**0.3218**	**0.0486**	**-0.1033**	**-0.1307**		-0.4086
32	**-0.0866**	**0.3404**	**0.0429**	-0.0179	**-0.1105**		-0.1495
33	**-0.0899**	**0.3644**	**0.0379**	0.0345	-0.1166		-0.0369
34	**-0.0802**	**0.3993**	**0.0366**	0.0468	-0.1225		-0.3564
35	**-0.0741**	**0.4259**	**0.0359**	0.0410	-0.1457		-0.5621
36	**-0.0725**	**0.4104**	**0.0361**	0.0257			-0.6189
37	**-0.0762**	**0.4474**	**0.0416**	-0.0314			-0.5162
38	**-0.0905**	**0.4715**	**0.0453**	-0.0809			0.0098
39	**-0.0990**	**0.4521**	**0.0444**	-0.0867			0.3427
40	**-0.1062**	**0.4707**	**0.0418**	-0.1204			0.6351
41	**-0.1299**	**0.4741**	**0.0416**	-0.1419			1.6004
42	**-0.1231**	**0.3462**	**0.0424**	-0.1272			1.3771
43	**-0.0757**	0.1881	**0.0495**	-0.0826			-0.6233
44	-0.0359	0.1113	**0.0599**	-0.0970			-2.4015
45	-0.0104	-0.0477	**0.0806**	-0.0547			**-3.6654**
46	0.0142	-0.1178	**0.0967**				**-4.8912**
47	-0.0149	-0.1106	**0.1082**				**-3.6658**
48	-0.0432	-0.0940	**0.1066**				-2.3195
49	**-0.0738**	-0.1186	**0.0996**				-0.7745
50	**-0.0903**	-0.0423	**0.0837**				0.1637
51	**-0.0863**	-0.0259	**0.0619**				0.0935
52	**-0.0860**	0.0766	0.0359				0.2574
53	**-0.1133**	0.1095	0.0175				1.7711
54	**-0.1305**	0.2696	0.0056				2.7245
55	-0.1141	0.2572	-0.0050				1.9071
56	-0.1018	0.2967	0.0001				1.2352
57	-0.0114	0.3009	0.0059				-3.9150
58	0.0047	0.4180	0.0114				-4.9085
59	-0.0534	0.3186	0.0047				-1.5017
60	-0.1370	0.4306	0.0080				3.3958

Note: Coefficients in **BOLD** are significant at the 90% level.

Figure 5: Residual Cohort Effects, Female First Marriage

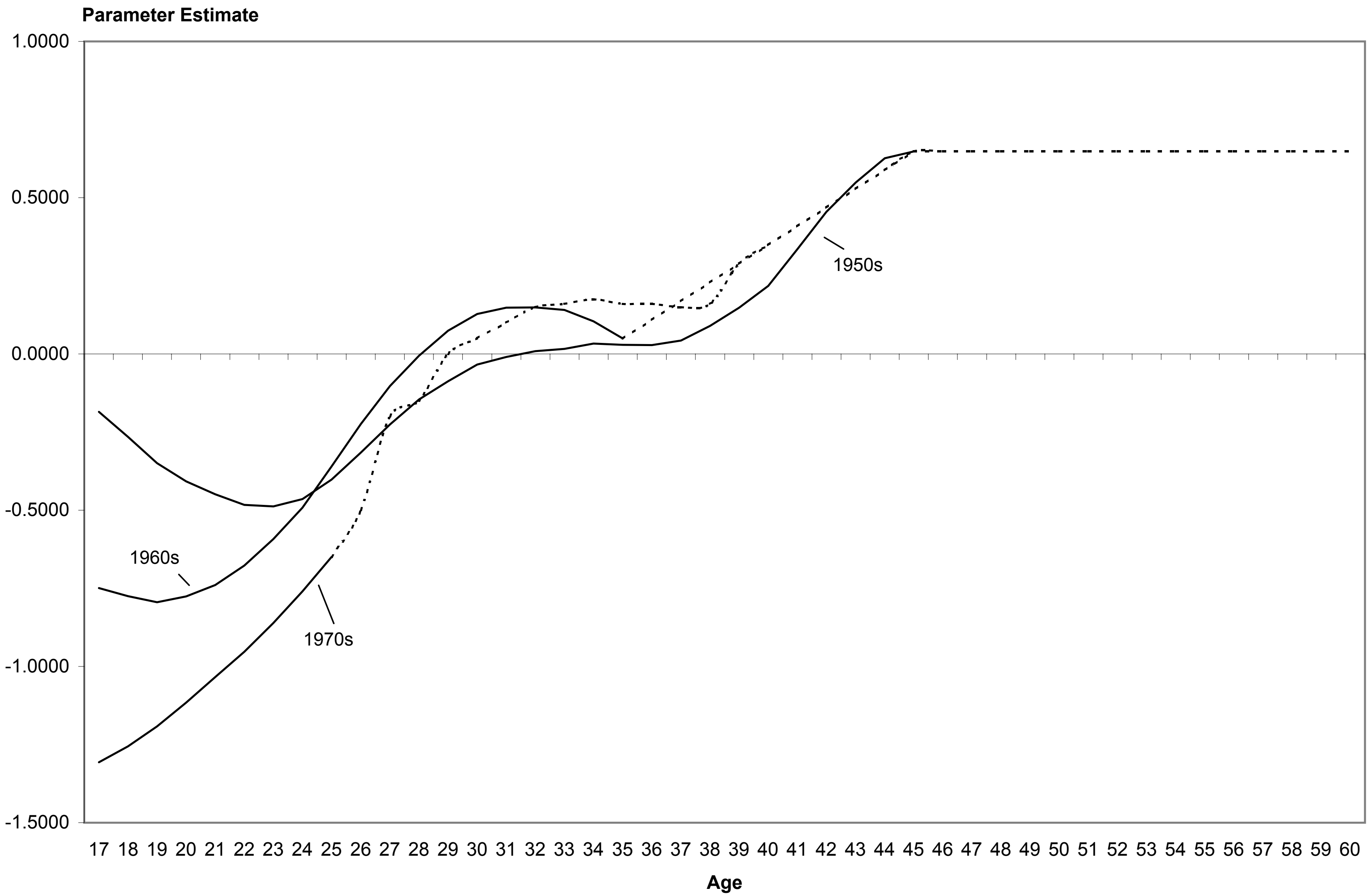


Figure 6: Residual Cohort Effects, Male First Marriage

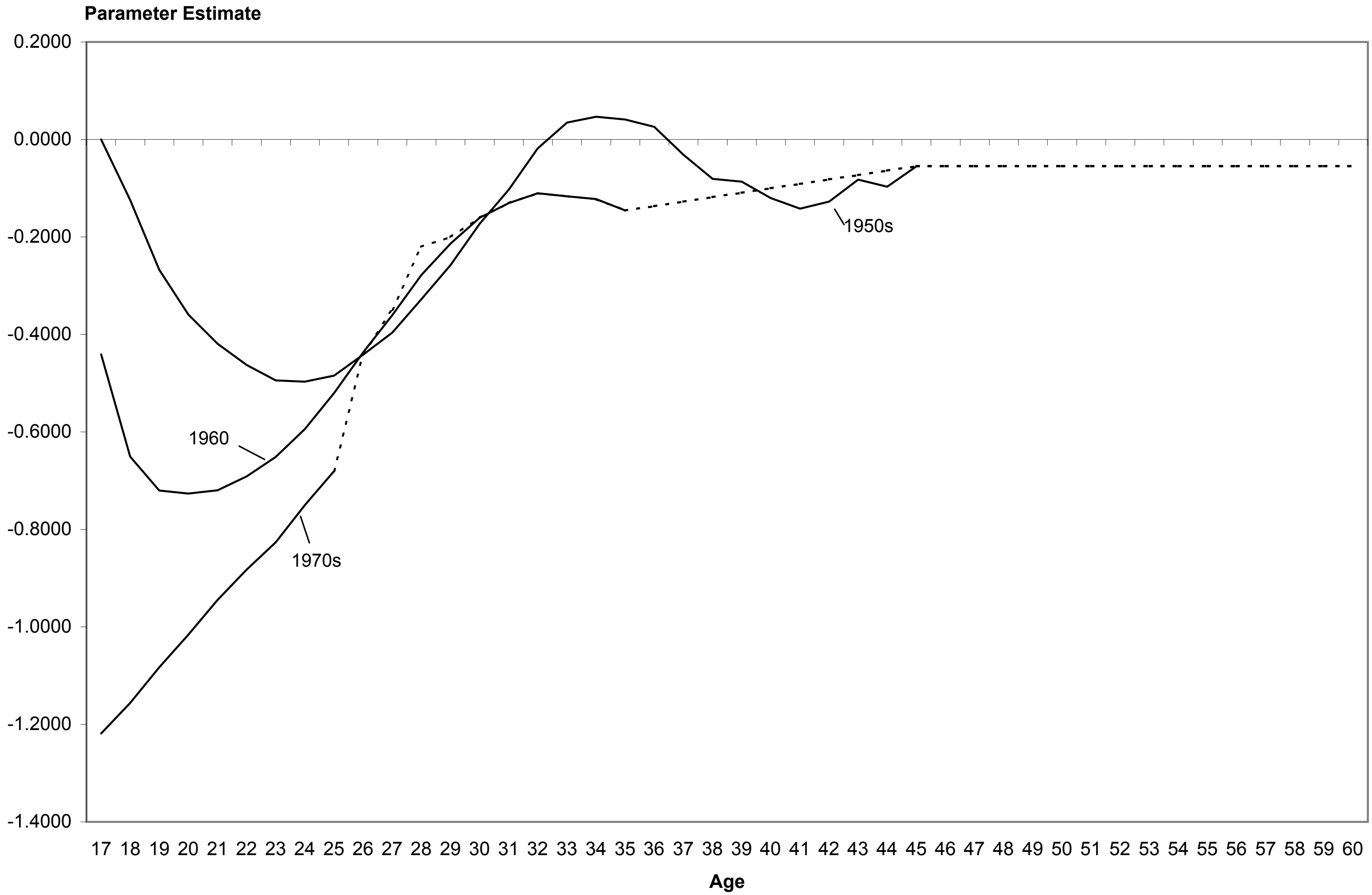


Table 4.
Potential Pairs Logistic Model Results

	Men's First Marriages	
Intercept	-3.953	*
	(0.070)	
Wife's marriage number	0.973	*
	(0.060)	
Wife has higher education	-1.147	*
	(0.043)	
Wife has lower education	-1.104	*
	(0.046)	
ALE quintile difference[1]	-0.062	*
	(0.008)	
ALE quintile difference squared[1]	-0.047	*
	(0.004)	
Age difference less than -6 spline	0.390	*
	(0.016)	
Age difference -6 to -1 spline	0.570	*
	(0.020)	
Age difference equal 0 dummy	-0.267	*
	(0.053)	
Age difference equal 1 dummy	-0.120	
	(0.051)	
Age difference 2 to 7 spline	-0.216	*
	(0.010)	
Age difference greater than 7 spline	-0.238	*
	(0.007)	
N	695,556	
-2 Log Likelihood	57,674	

Standard errors in parentheses, * = p < 0.01
1) Average Lifetime Earnings (ALE) quintile is calculated as husband's ALE quintile minus wife's ALE.

Figure 7: Distribution of Differences in Age Among Men's First Marriages
Simulated Marriages 2002-2076 and SIPP Marriages 1994-1996

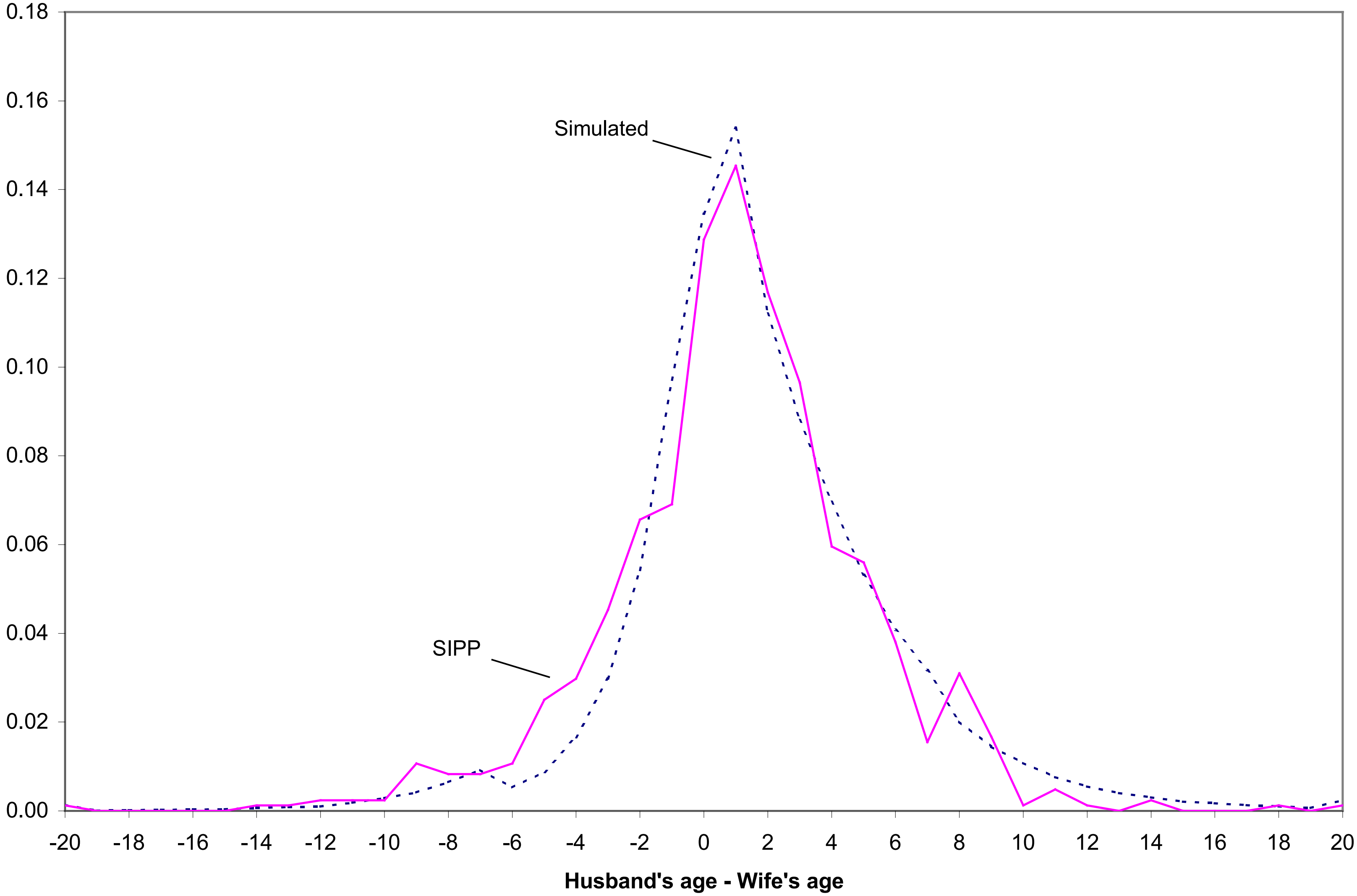


Figure 8: Distribution of Differences in Education Among Men's First Marriages, Simulated Marriages 2002-2076 and SIPP Marriages 1994-1996

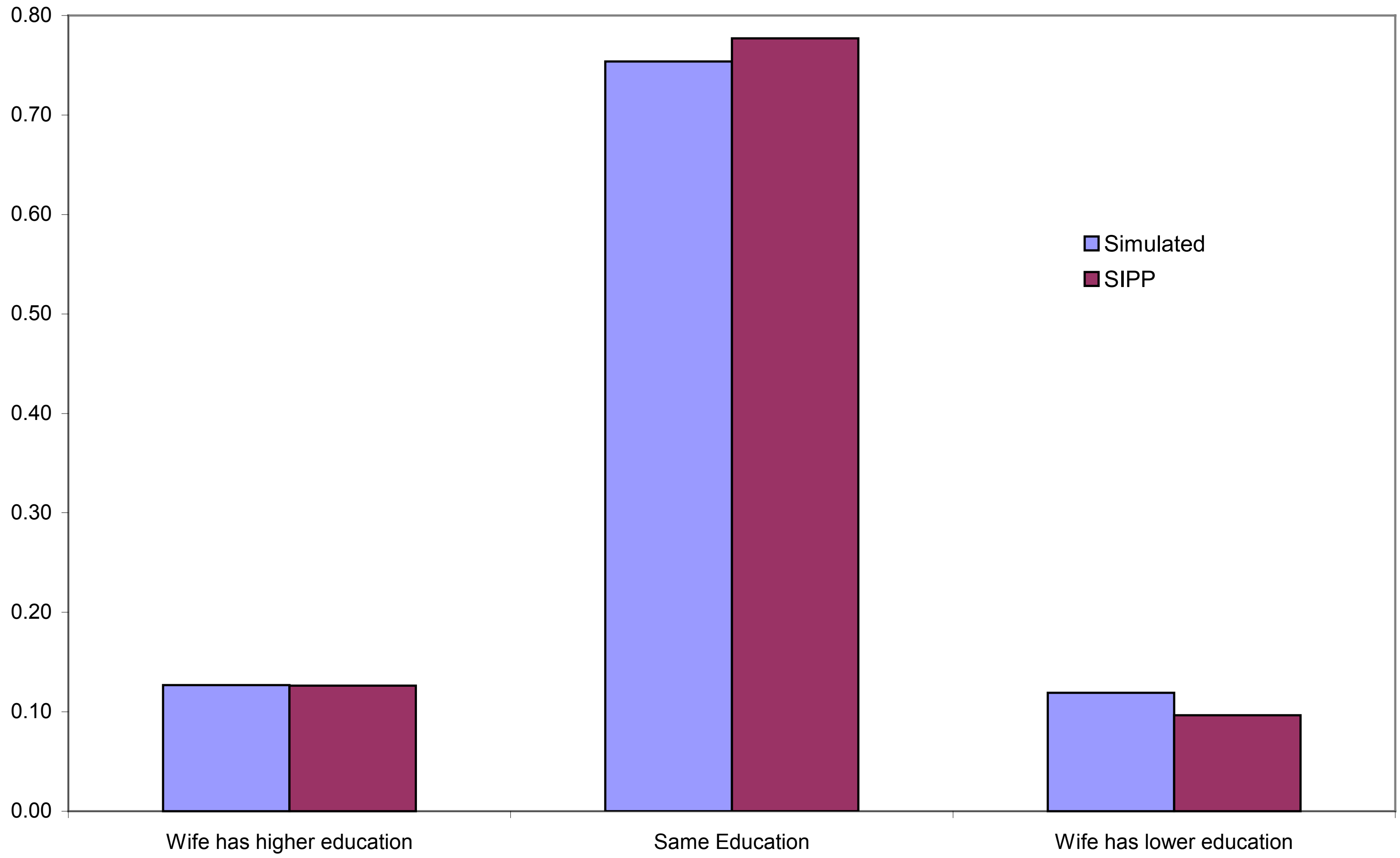


Figure 9: Distribution of Differences in Earnings Quintiles Among Men's First Marriages, Simulated Marriages 2002-2076 and SIPP Marriages 1994-1996

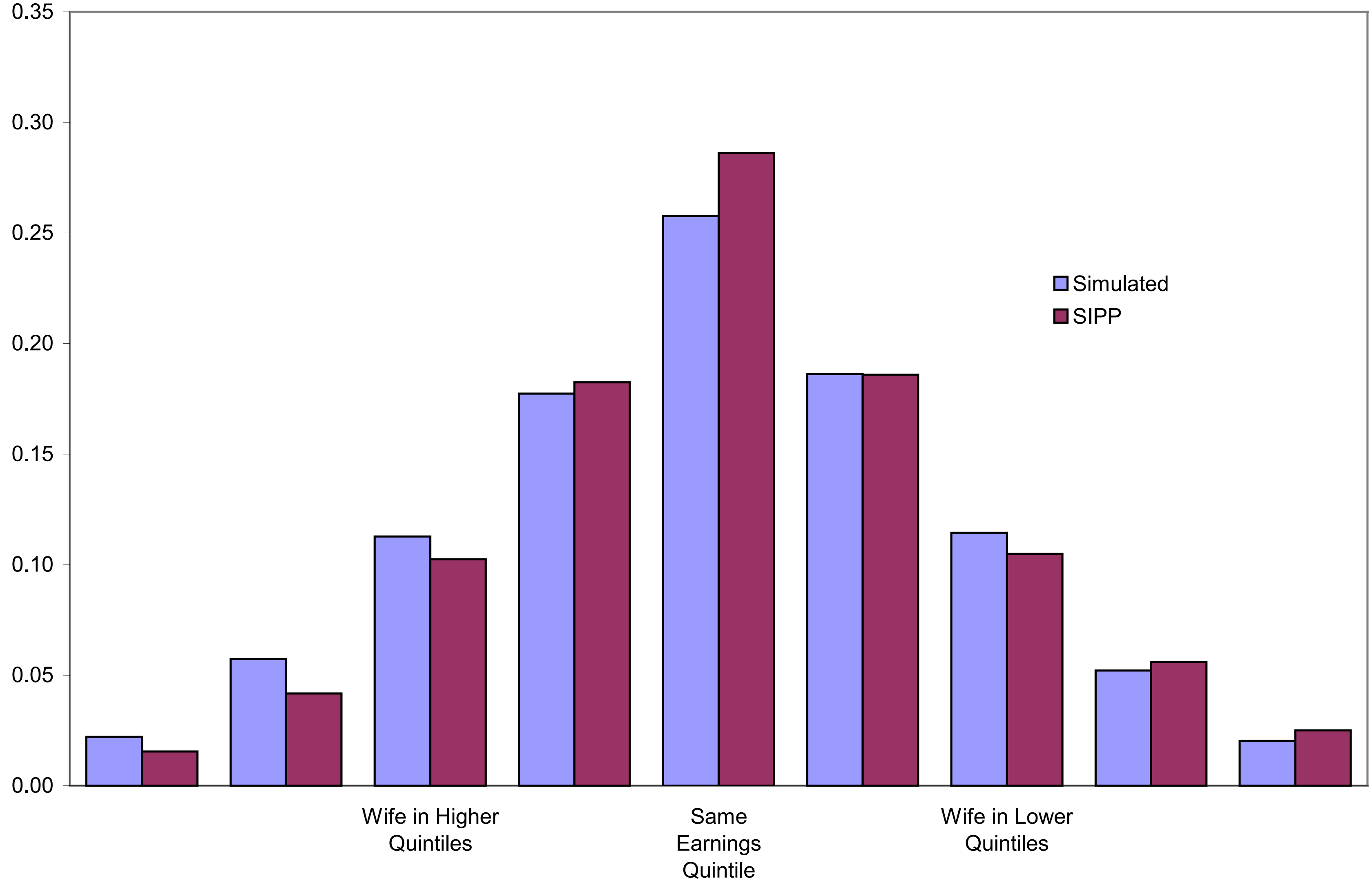
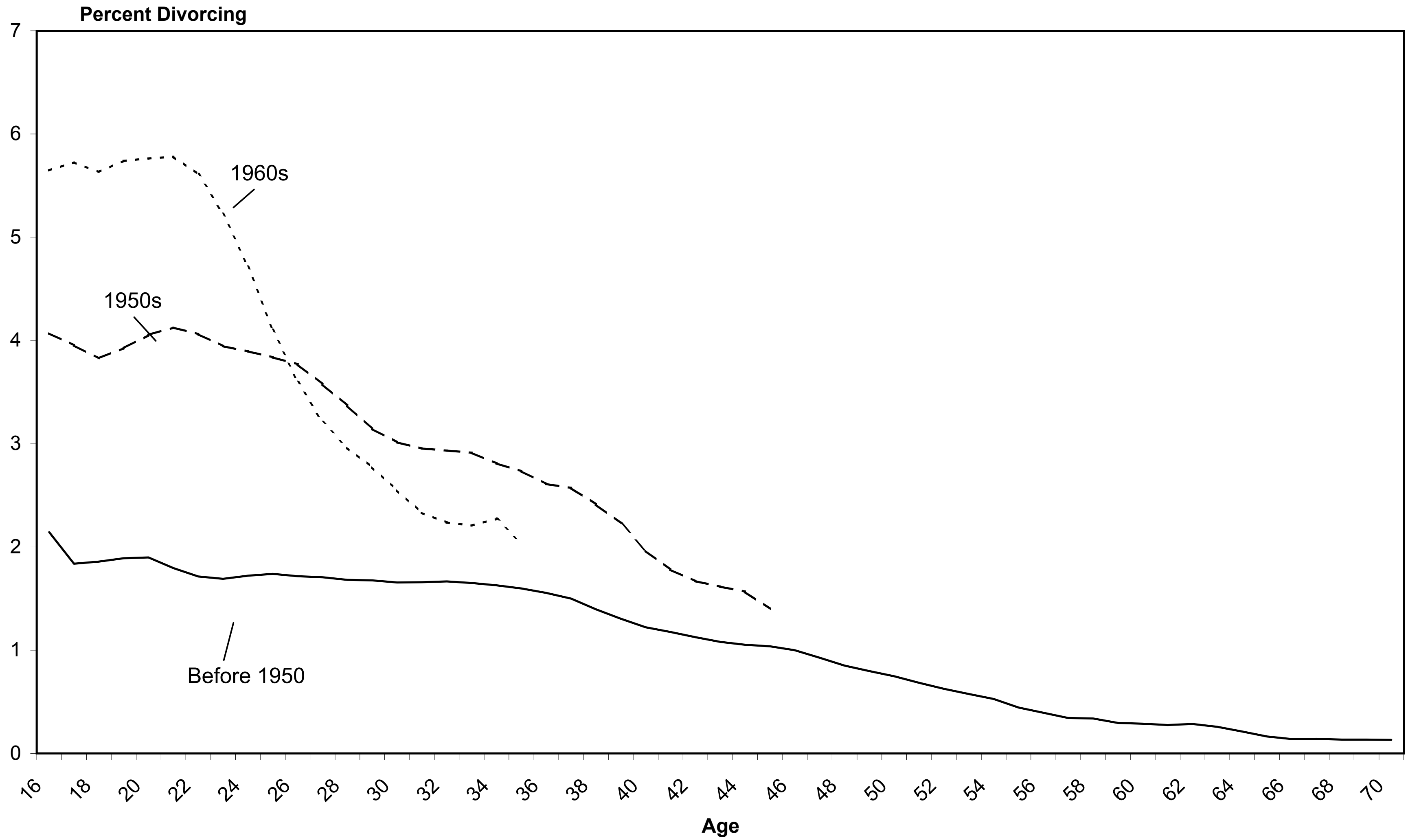
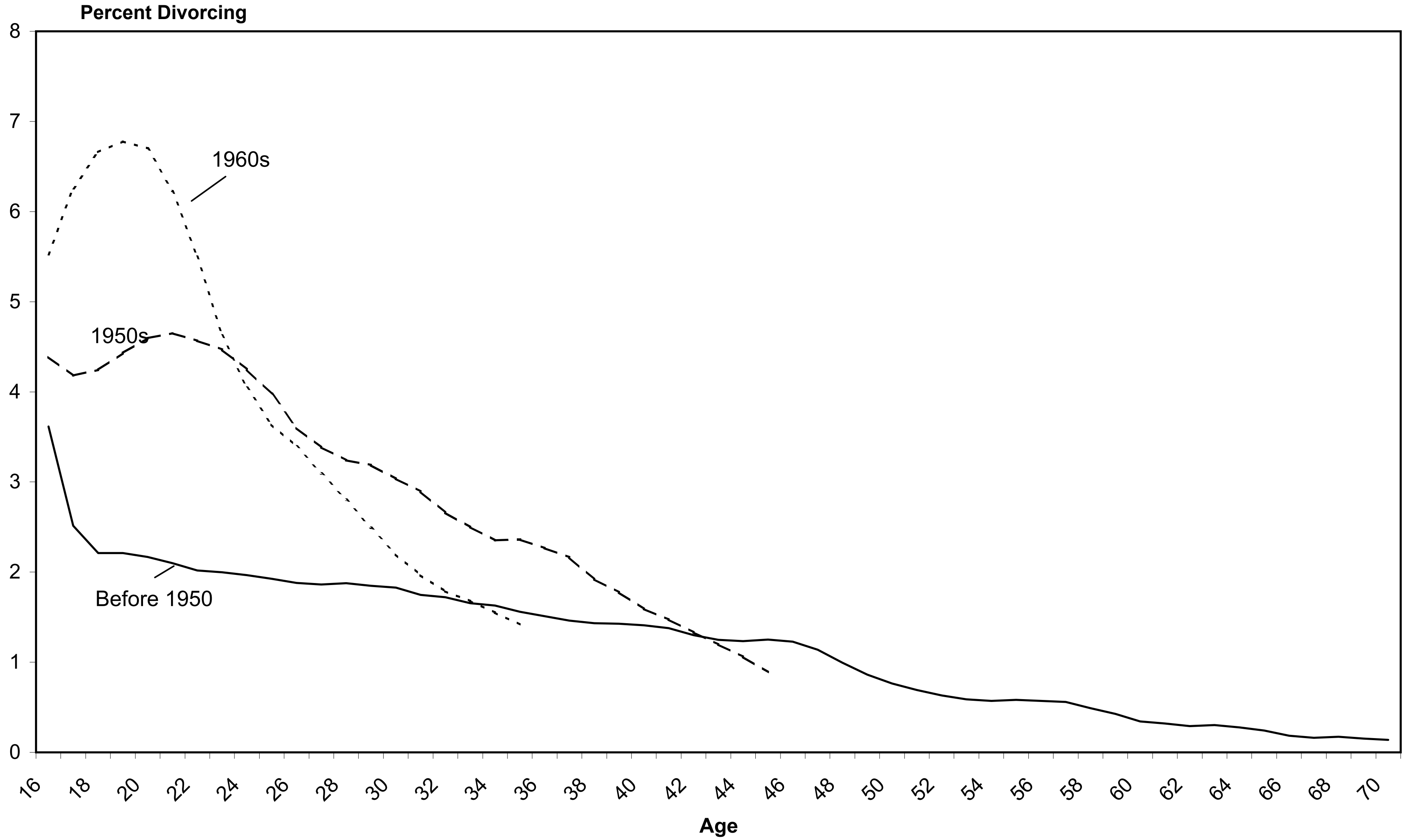


Figure 10: Female Divorce Rates by Cohort and Age



Source: Congressional Budget Office based on data from the 1996 Survey of Income and Program Participation.

Figure 11: Male Divorce Rates by Cohort and Age



Source: Congressional Budget Office based on the data from the 1996 Survey of Income and Program Participation.

Table 5.
Age-Centered Regression Coefficients, Female Divorce

Age	**Coefficient on:**								
Group	**Age**	**Educ**	**Duration**	**Durat Sq**	**Income**	**Prev Marr**	**Cohort50**	**Cohort60**	**Constant**
17	0.0543	0.2758	0.0000	0.0161	0.0604		0.4342	**0.9722**	-5.1091
18	-0.0097	0.1493	-0.1871	0.1004	**0.0704**		**0.6378**	**0.9239**	**-3.9443**
19	**-0.1305**	**0.1641**	**0.2359**	0.0066	**0.0897**		**0.7421**	**1.0766**	**-2.2157**
20	**-0.1777**	**0.1270**	**0.3546**	**-0.0145**	**0.0897**		**0.7757**	**1.1460**	**-1.4277**
21	**-0.1821**	0.1042	**0.3838**	**-0.0198**	**0.0899**		**0.7881**	**1.1581**	**-1.3668**
22	**-0.1757**	0.0708	**0.3707**	**-0.0198**	**0.0874**		**0.7996**	**1.1335**	**-1.4377**
23	**-0.1648**	0.0306	**0.3462**	**-0.0186**	**0.0848**		**0.7830**	**1.0674**	**-1.5608**
24	**-0.1498**	0.0016	**0.3161**	**-0.0172**	**0.0837**		**0.7604**	**0.9687**	**-1.7622**
25	**-0.1491**	0.0201	**0.3012**	**-0.0154**	**0.0859**	**0.6304**	**0.7155**	**0.8218**	**-1.7486**
26	**-0.1334**	0.0049	**0.2592**	**-0.0127**	**0.0863**	**0.6340**	**0.6721**	**0.6813**	**-1.9579**
27	**-0.1215**	-0.0114	**0.2108**	**-0.0096**	**0.0868**	**0.6064**	**0.6191**	**0.5448**	**-2.0599**
28	**-0.1048**	-0.0300	**0.1719**	**-0.0075**	**0.0882**	**0.6035**	**0.5699**	**0.4142**	**-2.3213**
29	**-0.0861**	-0.0505	**0.1336**	**-0.0055**	**0.0874**	**0.6136**	**0.5154**	**0.3073**	**-2.6473**
30	**-0.0694**	**-0.0846**	**0.1041**	**-0.0041**	**0.0871**	**0.6158**	**0.4697**	**0.2288**	**-2.9740**
31	**-0.0523**	**-0.1223**	**0.0840**	**-0.0032**	**0.0861**	**0.6164**	**0.4390**	**0.1573**	**-3.3570**
32	**-0.0443**	**-0.1410**	**0.0744**	**-0.0030**	**0.0865**	**0.6207**	**0.3988**	0.0999	**-3.5204**
33	**-0.0398**	**-0.1331**	**0.0554**	**-0.0021**	**0.0857**	**0.6185**	**0.3712**	0.0637	**-3.5659**
34	**-0.0390**	**-0.1116**	**0.0522**	**-0.0020**	**0.0863**	**0.6229**	**0.3536**	0.0199	**-3.5703**
35	**-0.0485**	-0.0593	**0.0420**	**-0.0016**	**0.0840**	**0.6180**	**0.3309**	-0.0434	**-3.1774**
36	**-0.0574**	0.0327	**0.0408**	**-0.0016**	**0.0830**	**0.5967**	**0.2954**		**-2.8440**
37	**-0.0650**	0.0796	**0.0338**	**-0.0014**	**0.0767**	**0.6298**	**0.2938**		**-2.5039**
38	**-0.0726**	**0.1074**	0.0254	-0.0011	**0.0725**	**0.6581**	**0.2656**		**-2.1389**
39	**-0.0761**	**0.1306**	0.0132	-0.0007	**0.0674**	**0.6709**	**0.2280**		**-1.8909**
40	**-0.0713**	**0.1437**	0.0075	-0.0006	**0.0637**	**0.6954**	**0.2096**		**-2.0240**
41	**-0.0629**	**0.1395**	-0.0004	-0.0003	**0.0621**	**0.6993**	**0.1824**		**-2.2789**
42	**-0.0495**	**0.1493**	-0.0070	0.0000	**0.0638**	**0.7239**	**0.1190**		**-2.8274**
43	**-0.0432**	**0.1678**	-0.0042	-0.0001	**0.0627**	**0.7327**	0.1014		**-3.1162**
44	**-0.0409**	**0.2085**	-0.0008	-0.0003	**0.0621**	**0.7434**	0.0841		**-3.2389**
45	**-0.0482**	**0.2429**	0.0002	-0.0004	**0.0637**	**0.7490**	-0.0144		**-2.9070**
46	**-0.0444**	**0.2635**	0.0116	-0.0008	**0.0657**	**0.7489**			**-3.1332**
47	**-0.0434**	**0.2572**	0.0120	-0.0009	**0.0605**	**0.7521**			**-3.0950**
48	**-0.0532**	**0.2318**	0.0109	-0.0010	**0.0593**	**0.7411**			**-2.5465**
49	**-0.0532**	0.1314	0.0173	**-0.0013**	**0.0618**	**0.7368**			**-2.5271**
50	**-0.0380**	0.0346	0.0281	**-0.0017**	**0.0614**	**0.7013**			**-3.2604**
51	-0.0361	-0.0345	0.0346	**-0.0020**	**0.0619**	**0.6379**			**-3.3134**
52	-0.0378	-0.0838	**0.0491**	**-0.0024**	**0.0676**	**0.5849**			**-3.2983**
53	**-0.0458**	-0.1341	**0.0651**	**-0.0028**	**0.0701**	**0.5816**			**-3.0124**
54	**-0.0452**	-0.0628	**0.0671**	**-0.0028**	**0.0663**	**0.5461**			**-3.0347**
55	**-0.0561**	0.0207	**0.0552**	**-0.0025**	**0.0600**	**0.5374**			-2.3724
56	**-0.0573**	0.1384	0.0383	**-0.0020**	**0.0581**	**0.5230**			-2.1941
57	**-0.0578**	0.2248	0.0137	**-0.0014**	**0.0482**	**0.4981**			-1.9207
58	-0.0377	0.3774	-0.0189	-0.0008	**0.0353**	**0.4333**			-2.6780
59	-0.0024	**0.4647**	-0.0297	-0.0006	**0.0341**	**0.4387**			**-4.6634**
60	0.0097	**0.4895**	-0.0185	-0.0009	**0.0422**	**0.4158**			**-5.4251**
61	0.0035	**0.5604**	0.0014	**-0.0014**	**0.0405**	**0.4452**			**-5.2106**
62	-0.0120	**0.6306**	0.0363	**-0.0023**	**0.0461**	0.4314			-4.5039
63	-0.0098	**0.7281**	**0.0748**	**-0.0031**	**0.0631**	0.4435			-5.0705
64	-0.0727	**0.7160**	**0.0832**	**-0.0033**	**0.0611**	0.3757			-1.0626
65	-0.0958	**0.6852**	**0.0845**	**-0.0033**	0.0575	0.4366			0.3086
66	-0.0994	0.5004	0.0579	**-0.0026**	**0.0657**	0.4600			0.7356
67	-0.0379	0.2292	0.0042	-0.0012	**0.0912**	**0.6709**			-3.2858
68	0.0464	-0.1925	-0.0345	-0.0002	**0.1144**	0.8795			-9.0234
69	0.0862	0.1137	-0.0368	0.0000	**0.1551**	1.0658			-12.2533
70	0.0605	0.3973	-0.0539	0.0004	**0.2015**	1.0900			-10.5904

Note: Coefficients in **BOLD** are significant at the 90 percent level.

Table 6.
Age-Centered Regression Coefficients, Male Divorce

Age	**Coefficient on:**								
Group	**Age**	**Educ.**	**Duration**	**Durat Sq**	**Income**	**Prev Marr**	**Cohort50**	**Cohort60**	**Constant**
17	**-1.1723**	-1.6886	0.1394	0.1394	0.1472		0.8718	1.1750	15.5299
18	-0.1229	0.1031	**-1.0627**	**0.3932**	**0.0781**		**0.6675**	**1.1775**	-1.6450
19	**-0.1372**	**0.2257**	-0.0149	**0.0898**	**0.0383**		**0.6617**	**1.1010**	**-1.6426**
20	**-0.2106**	**0.1728**	**0.2063**	**0.0271**	**0.0201**		**0.7039**	**1.0119**	-0.0718
21	**-0.2233**	0.1190	**0.2824**	0.0090	0.0154		**0.7164**	**0.9764**	0.2337
22	**-0.2221**	0.0498	**0.3217**	-0.0028	0.0111		**0.7339**	**0.9544**	0.2774
23	**-0.2164**	-0.0224	**0.3470**	**-0.0103**	0.0065		**0.7311**	**0.8917**	0.2459
24	**-0.2004**	**-0.1067**	**0.3411**	**-0.0131**	0.0022		**0.7155**	**0.8311**	0.0256
25	**-0.1886**	**-0.1449**	**0.3399**	**-0.0146**	-0.0017	**0.5890**	**0.6988**	**0.7637**	-0.1720
26	**-0.1657**	**-0.1796**	**0.3101**	**-0.0136**	-0.0082	**0.6695**	**0.6650**	**0.6845**	**-0.5560**
27	**-0.1473**	**-0.1910**	**0.2883**	**-0.0128**	**-0.0142**	**0.7271**	**0.6302**	**0.5915**	**-0.8715**
28	**-0.1324**	**-0.1951**	**0.2612**	**-0.0113**	**-0.0181**	**0.7615**	**0.6056**	**0.5070**	**-1.1242**
29	**-0.1174**	**-0.1814**	**0.2322**	**-0.0098**	**-0.0210**	**0.7695**	**0.5883**	**0.4096**	**-1.3952**
30	**-0.1074**	**-0.1824**	**0.2012**	**-0.0082**	**-0.0231**	**0.7497**	**0.5443**	**0.2995**	**-1.5148**
31	**-0.0967**	**-0.1944**	**0.1688**	**-0.0067**	**-0.0222**	**0.6919**	**0.5297**	**0.2096**	**-1.6719**
32	**-0.0838**	**-0.2185**	**0.1399**	**-0.0055**	**-0.0201**	**0.6283**	**0.5052**	0.1259	**-1.9149**
33	**-0.0705**	**-0.2334**	**0.1071**	**-0.0041**	**-0.0191**	**0.5853**	**0.4727**	0.0527	**-2.1666**
34	**-0.0611**	**-0.2592**	**0.0861**	**-0.0034**	**-0.0172**	**0.5722**	**0.4298**	-0.0044	**-2.3554**
35	**-0.0523**	**-0.2633**	**0.0700**	**-0.0028**	**-0.0150**	**0.5757**	**0.4151**	-0.0512	**-2.5707**
36	**-0.0492**	**-0.2357**	**0.0658**	**-0.0027**	**-0.0145**	**0.5836**	**0.3743**		**-2.6462**
37	**-0.0453**	**-0.2078**	**0.0551**	**-0.0023**	**-0.0169**	**0.6282**	**0.3196**		**-2.7123**
38	**-0.0426**	**-0.1799**	**0.0476**	**-0.0020**	**-0.0189**	**0.6404**	**0.2534**		**-2.7369**
39	**-0.0454**	**-0.1269**	**0.0386**	**-0.0017**	**-0.0227**	**0.6350**	**0.1962**		**-2.5360**
40	**-0.0442**	-0.0915	**0.0383**	**-0.0017**	**-0.0262**	**0.6566**	0.0997		**-2.5428**
41	**-0.0414**	-0.0352	0.0314	**-0.0015**	**-0.0296**	**0.6637**	0.0051		**-2.5746**
42	**-0.0394**	0.0052	0.0244	-0.0013	**-0.0314**	**0.6574**	-0.0674		**-2.5969**
43	**-0.0326**	0.0604	0.0175	-0.0011	**-0.0326**	**0.6647**	-0.1432		**-2.8334**
44	-0.0204	0.0874	0.0202	**-0.0013**	**-0.0314**	**0.6681**	**-0.2542**		**-3.3595**
45	-0.0264	**0.1461**	0.0139	-0.0012	**-0.0307**	**0.6392**	**-0.3389**		**-3.0313**
46	**-0.0346**	**0.1865**	0.0074	-0.0010	**-0.0275**	**0.5919**			**-2.6187**
47	**-0.0566**	**0.1652**	0.0037	-0.0011	**-0.0265**	**0.5580**			**-1.5193**
48	**-0.0852**	0.0948	0.0039	**-0.0012**	**-0.0285**	**0.5429**			-0.0944
49	**-0.1000**	0.0356	-0.0045	-0.0010	**-0.0324**	**0.5155**			0.7462
50	**-0.0899**	-0.0500	-0.0066	-0.0010	**-0.0383**	**0.4857**			0.3825
51	**-0.0648**	-0.1122	-0.0007	**-0.0012**	**-0.0467**	**0.4545**			-0.7630
52	-0.0390	-0.1313	0.0036	-0.0014	**-0.0527**	**0.4397**			-2.0632
53	-0.0032	-0.0422	0.0063	-0.0015	**-0.0524**	**0.3654**			**-3.8923**
54	0.0109	0.0588	0.0148	**-0.0018**	**-0.0516**	**0.3155**			**-4.6891**
55	0.0163	0.0940	0.0277	**-0.0021**	**-0.0499**	**0.2724**			**-5.0848**
56	0.0067	0.1473	0.0347	**-0.0022**	**-0.0448**	**0.2560**			**-4.6894**
57	-0.0215	0.1545	0.0482	**-0.0025**	**-0.0413**	**0.2417**			**-3.2256**
58	**-0.0655**	0.0909	**0.0625**	**-0.0027**	**-0.0395**	**0.2495**			-0.9003
59	**-0.0745**	-0.0350	**0.0861**	**-0.0033**	**-0.0368**	0.2233			-0.5297
60	**-0.0760**	-0.0742	**0.0876**	**-0.0032**	-0.0302	0.2632			-0.5605
61	-0.0587	-0.1367	**0.0876**	**-0.0031**	-0.0269	0.2939			-1.6836
62	-0.0252	-0.0746	**0.0789**	**-0.0028**	-0.0233	0.3129			-3.8189
63	-0.0451	-0.0093	0.0630	**-0.0024**	**-0.0257**	0.3749			-2.5349
64	-0.0660	0.1021	0.0286	-0.0014	-0.0245	**0.5238**			-1.2080
65	-0.1109	0.1424	0.0196	-0.0012	-0.0334	**0.5892**			1.6943
66	**-0.1260**	0.2046	0.0298	-0.0014	-0.0319	**0.6737**			2.4686
67	**-0.1419**	0.1018	0.0324	-0.0014	-0.0249	**0.8168**			3.3959
68	-0.0786	0.1695	0.0278	-0.0013	-0.0053	**0.9273**			-0.9322
69	-0.0690	0.1962	0.0364	-0.0015	0.0262	**1.0116**			-1.8784
70	-0.0489	0.2294	0.0282	-0.0013	0.0802	**1.0093**			-3.4575

Note: Coefficients in **BOLD** are significant at the 90 percent level.

Figure 12: Residual Cohort Effects, Female Divorce

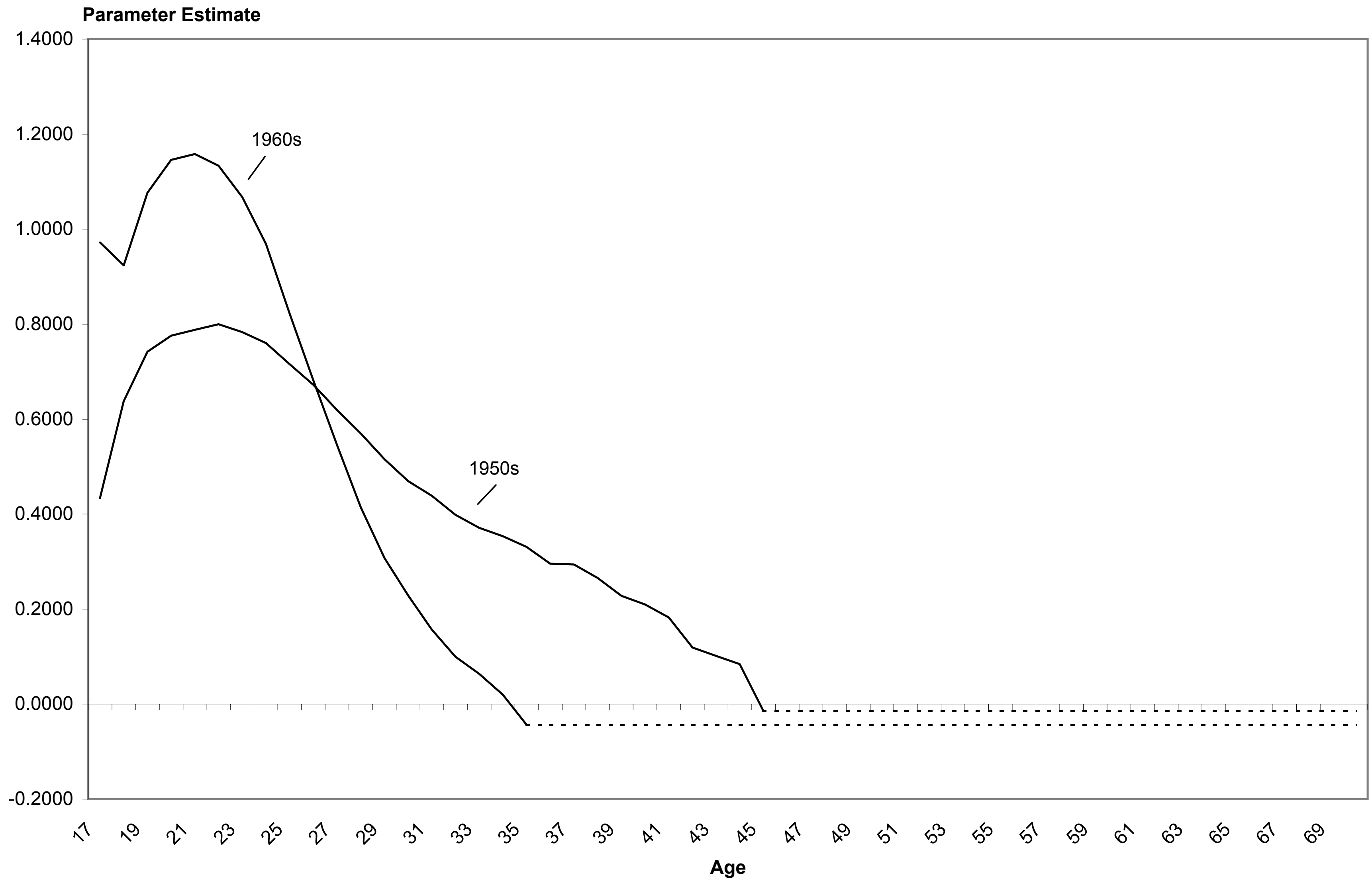


Figure 13: Residual Cohort Effects, Male Divorce

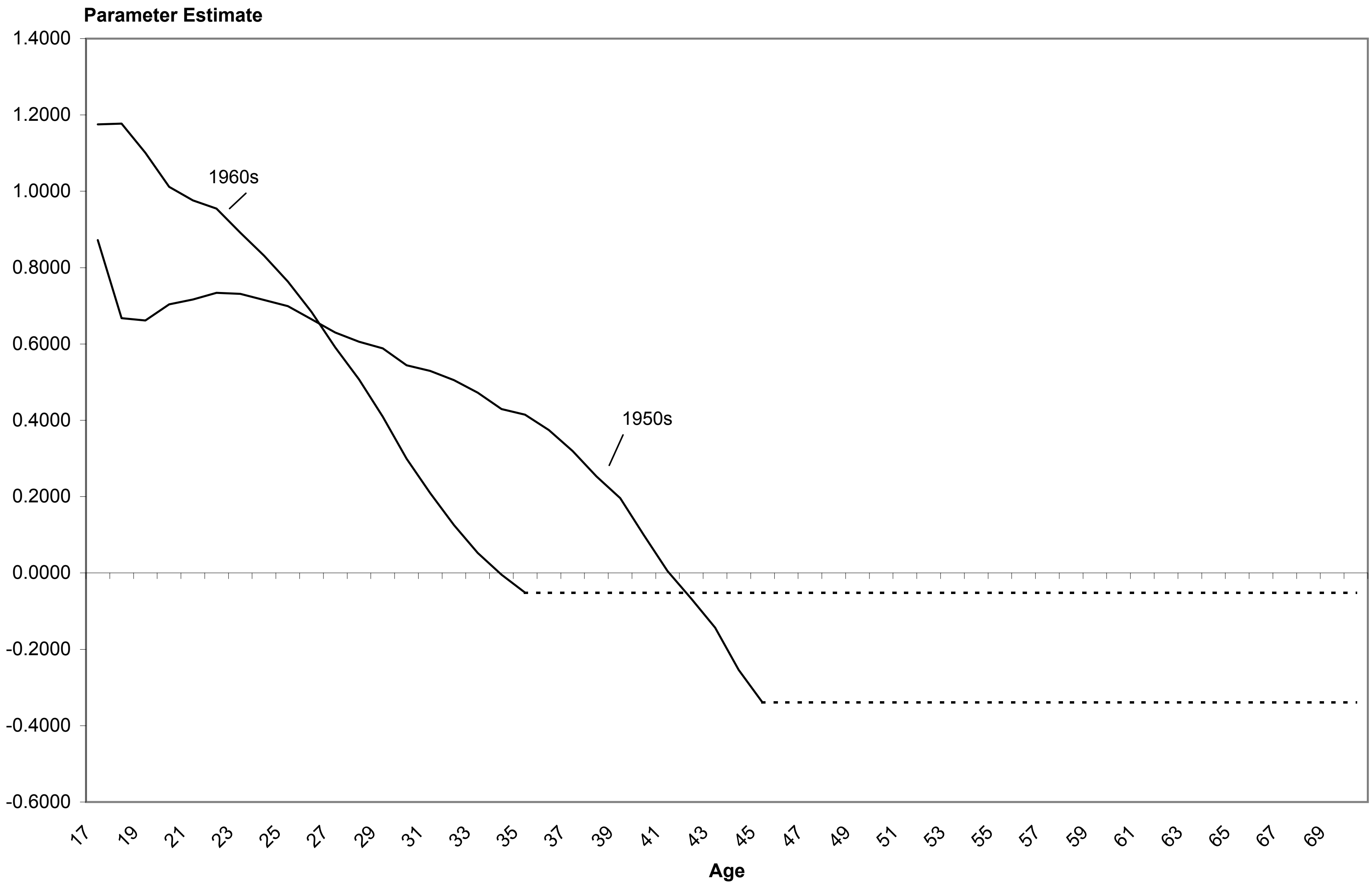


Figure 14: Historical Simulation Calibration Factors, Females

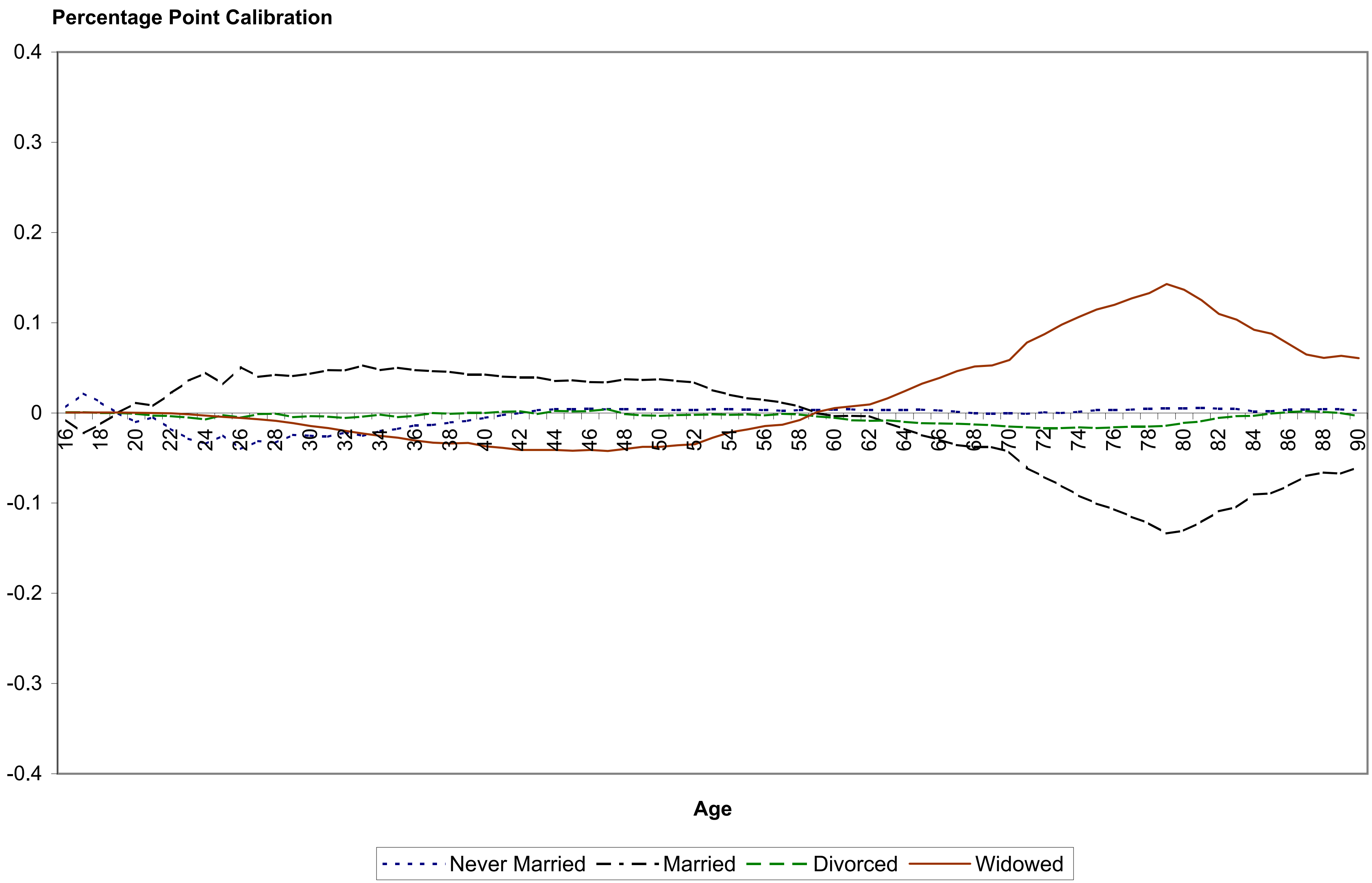


Figure 15: Historical Simulation Calibration Factors, Males

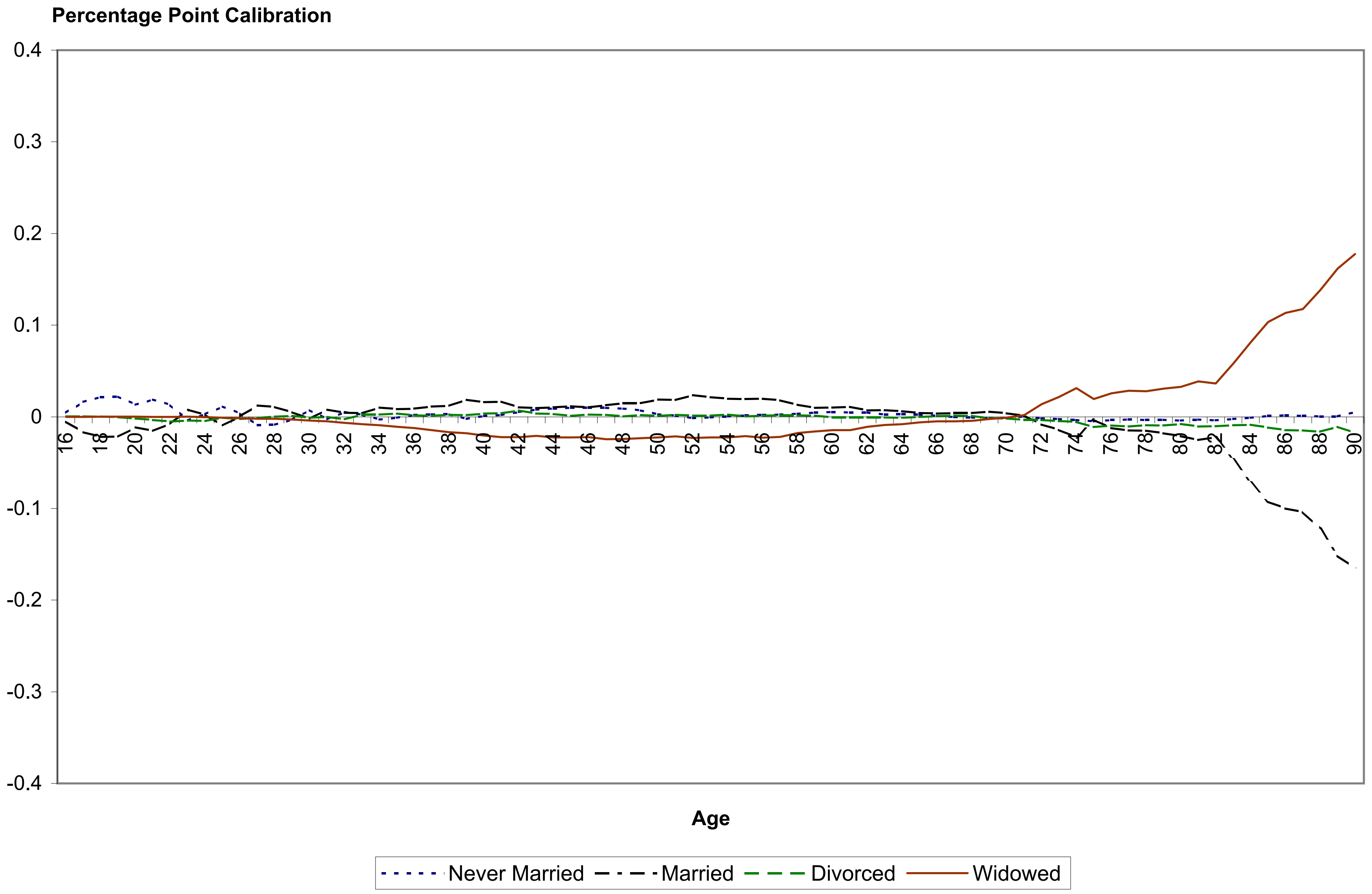


Figure 16: Marital Status for Females Ages 62-67, by Year

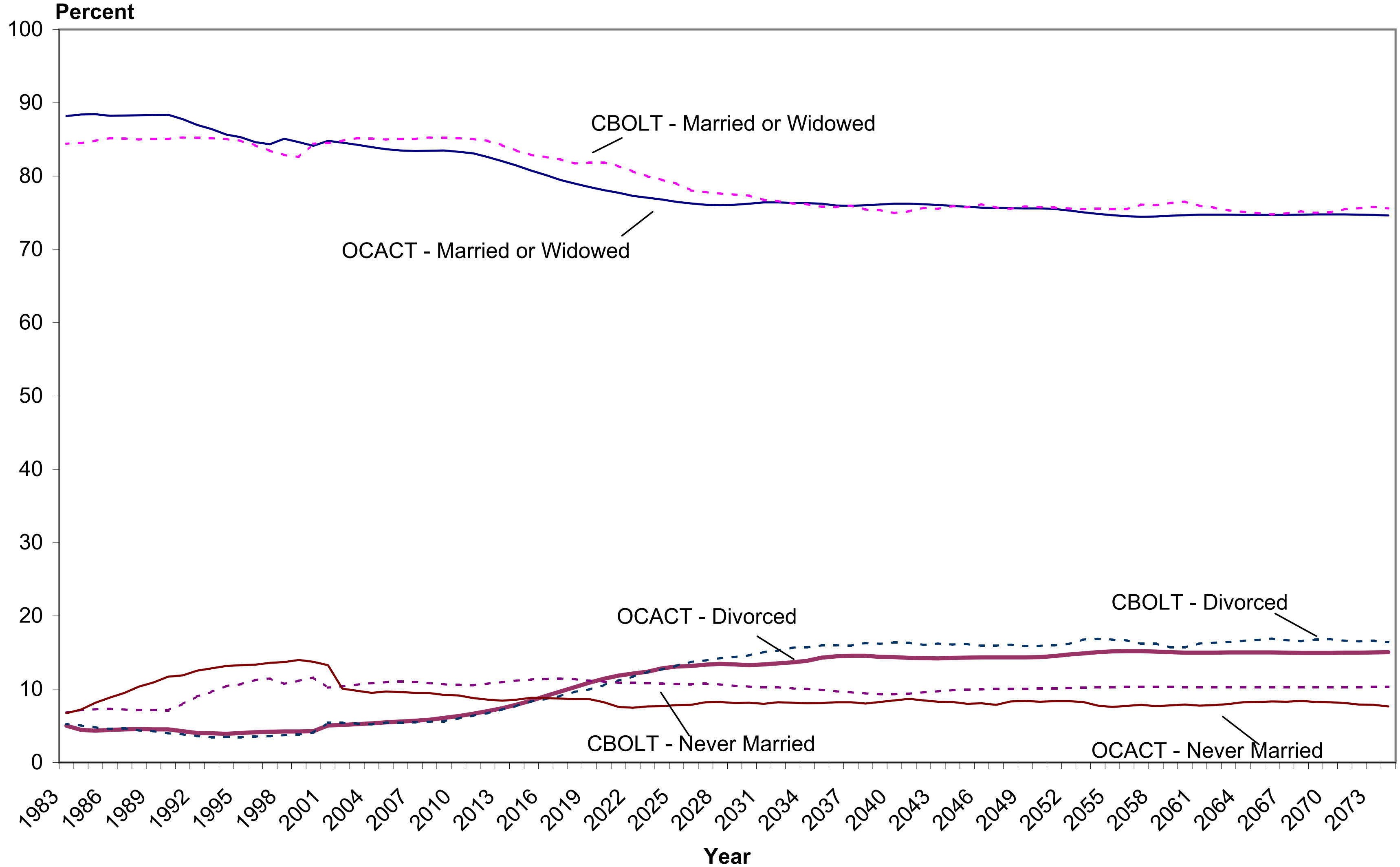


Figure 17: Marital Status for Males Ages 62-67, by Year

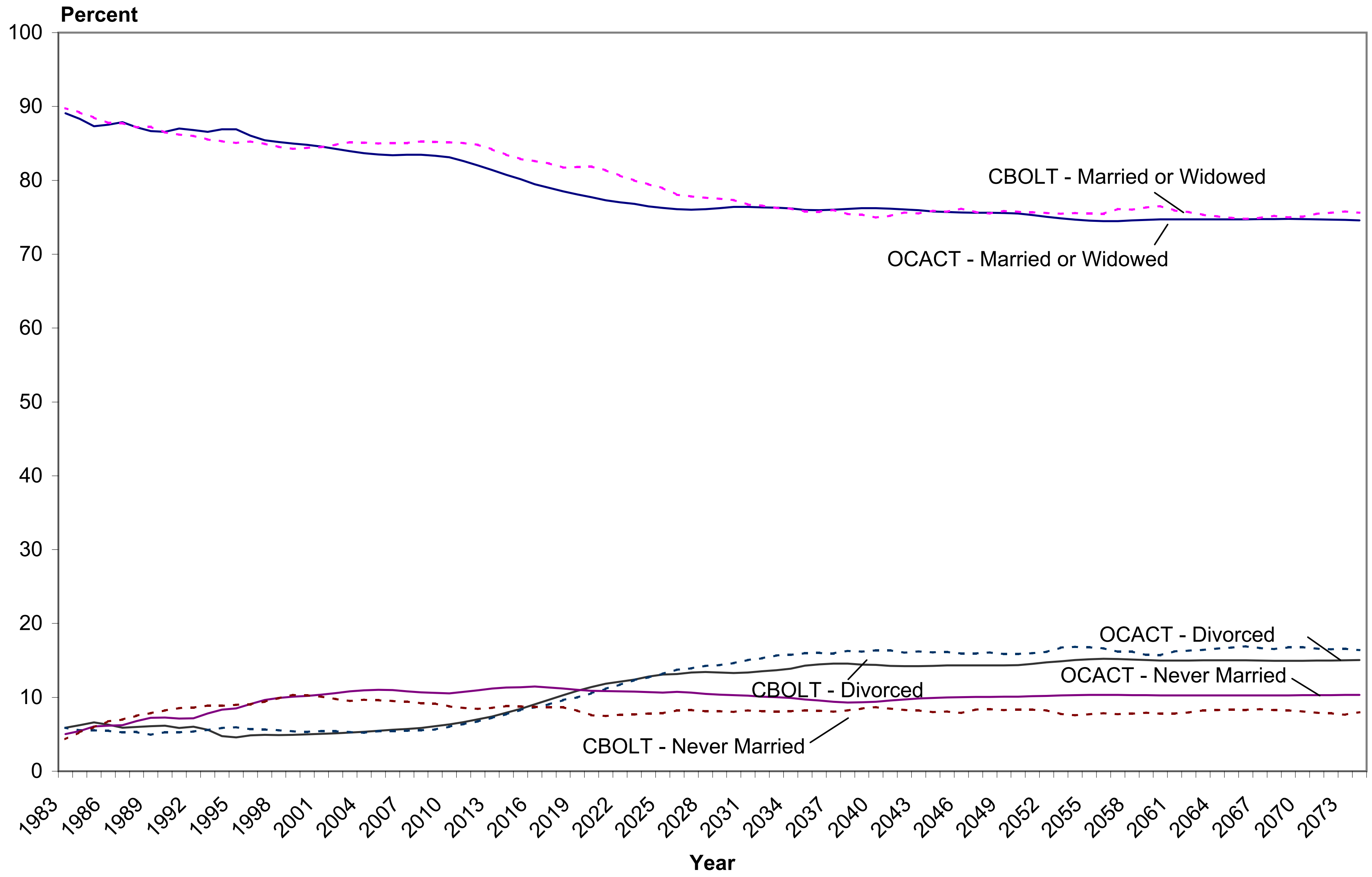


Table 7: Marital Status Distributions Over Time, by Sex and Year

	Social Security Administration (OCACT)				Congressional Budget Office Long-Term (CBOLT) Model			
	Never Married	**Married**	**Divorced**	**Widowed**	**Never Married**	**Married**	**Divorced**	**Widowed**
Overall Population								
2001	43%	44%	8%	5%	44%	44%	7%	6%
2025	41%	45%	9%	5%	42%	46%	6%	6%
2050	40%	45%	9%	6%	41%	46%	6%	7%
2075	40%	46%	9%	6%	40%	47%	6%	7%
Female Population								
2001	40%	43%	8%	8%	41%	43%	8%	9%
2025	38%	45%	10%	8%	38%	46%	7%	9%
2050	37%	45%	10%	9%	36%	47%	7%	10%
2075	36%	46%	10%	8%	35%	48%	7%	10%
Male Population								
2001	47%	45%	7%	2%	47%	45%	6%	3%
2025	45%	46%	7%	2%	46%	46%	5%	3%
2050	44%	46%	7%	3%	46%	45%	5%	4%
2075	43%	46%	7%	3%	45%	46%	5%	4%

Source: Authors' calculations based on SSA Area Population Projections 1941 to 2080, Office of the Chief Actuary.

Figure 18: Mean Age At First Marriage

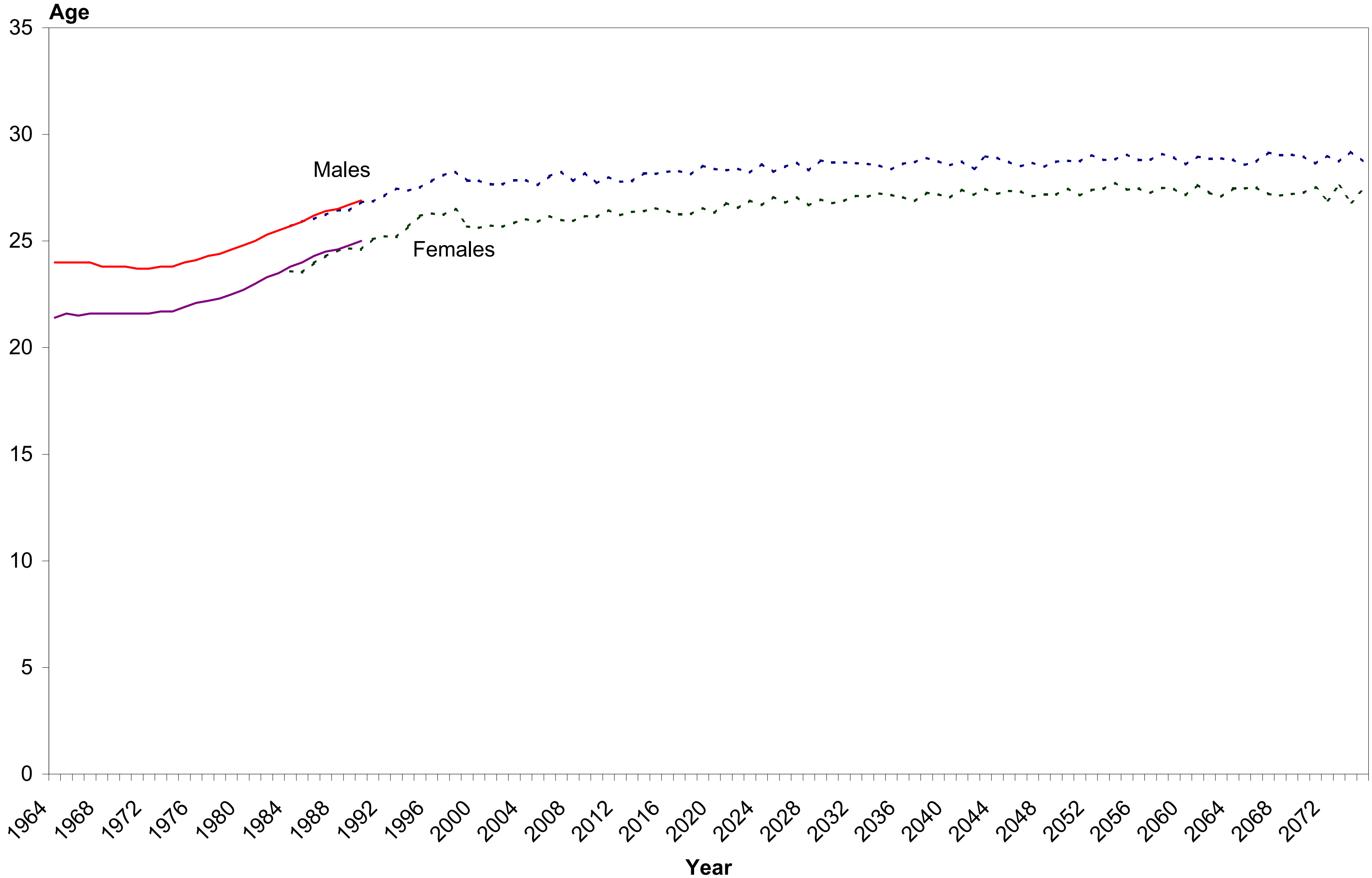


Table 8.
Distribution of Marital Durations at Time of Divorce, by Number of Years Married

	0	1 to 4	5 to 9	10 to14	15 or More
NCHS - Actual					
1990	3%	32%	28%	15%	22%
Projected					
2005	9%	30%	24%	16%	23%
2015	8%	28%	24%	15%	27%
2025	6%	31%	27%	14%	23%
2035	6%	30%	25%	15%	24%
2045	7%	28%	28%	13%	25%
2055	6%	30%	25%	14%	23%
2065	8%	31%	22%	14%	26%
2075	7%	29%	26%	13%	26%

Table 9.
Percentage of Nonwidowed Women Ages 62-67 Eligible for OAI Spouse Benefits, by Lifetime Earnings Deciles

	Year							
	2005	**2015**	**2025**	**2035**	**2045**	**2055**	**2065**	**2075**
Earnings Decile								
1	95%	88%	88%	86%	83%	86%	85%	87%
2	95%	89%	87%	87%	82%	87%	88%	89%
3	90%	89%	85%	89%	84%	88%	89%	87%
4	90%	83%	83%	86%	83%	85%	85%	87%
5	86%	80%	83%	85%	79%	84%	85%	85%
6	88%	83%	83%	82%	80%	82%	84%	83%
7	80%	81%	82%	79%	74%	80%	83%	81%
8	84%	80%	77%	78%	74%	77%	78%	77%
9	81%	77%	75%	71%	69%	76%	77%	78%
10	77%	75%	72%	64%	65%	68%	70%	72%
All	87%	83%	81%	81%	77%	81%	82%	83%