Cohort Size and The Marriage Market:

Explaining Nearly a Century of Changes in U.S. Marriage Rates*

Mary Ann Bronson† Maurizio Mazzocco‡

Current Draft, October, 2012. First draft, February, 2012

Abstract

We propose an explanation for almost a century of change in U.S. marriage rates. We do this in three stages. In the first stage, we use reduced-form methods to provide evidence on the following two results. First, a single variable, cohort size, can account for almost all the variation in marriage rates since the early 1930s for both blacks and whites. Second, an increase in cohort size reduces marriage rates, whereas a decline in cohort size has the opposite effect. The most convincing evidence on the relationship between cohort size and marriage rates is obtained by using as a source of exogenous variation differences in state mobilization rates during World War II. In the second stage, we develop a dynamic search model of the marriage market that can qualitatively generate the observed negative relationship between cohort size and marriage rates. We also derive a testable implication from the model: an increase in cohort size reduces spouses' age difference at marriage, and a decrease has the opposite effect. Lastly, we investigate whether the model is consistent with the data. We first use the derived implication to test the model and fail to reject it. We then estimate the model and evaluate whether it can quantitatively match the documented variation in marriage rates. We find that the search model can explain most of the observed variation.

*We are grateful to Guillermo Beylis, Pierre-André Chiappori, Edgar Cortes, Kenneth Mirkin, Dmitry Plotnikov, and Nico Voigtlaender for helpful comments. Research reported in this paper was supported by the National Institutes of Health under Award Numbers 2T32HD7545-11 and 5T32HD7545-10. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

†UCLA, Department of Economics, Bunche Hall, Los Angeles, CA. Email: mary.ann.bronson@gmail.com.

‡UCLA, Department of Economics, Bunche Hall, Los Angeles, CA. Email: mmazzocc@econ.ucla.edu.

1 Introduction

Changes in marital patterns have major implications for many variables of interest to economists and policy makers. They affect, to name a few, fertility rates, children's welfare, children's education, labor force participation, hours of work, income inequality, the fraction of individuals on welfare, population growth, and workers' productivity. In spite of this, no prevailing explanation accounts for the variation in marriage formation over time and across races. Existing theories, discussed in the next section, have either been empirically rejected, or apply only to specific periods or specific groups of individuals.

The main contribution of this paper is to provide an explanation for nearly a century of changes in U.S. marriage rates. We show that changes in cohort size on their own explain virtually the entire variation in marriage rates since the early 1930s for both blacks and whites. The paper consists of three parts.

In the first part, we present reduced-form evidence which indicates that an increase in cohort size generates a decline in marriage rates and that a reduction in cohort size has the opposite effect. We provide reduced-form evidence in three steps. We first employ time-series variation to study the relationship between changes in marriage rates and changes in cohort size. Our results show that, for both blacks and whites, there is a strong and negative relationship between these two variables. They also indicate that changes in cohort size account for around 70 percent of the variation in marriage rates for both black and white populations. We then use cross-state variation to provide additional evidence on the link between cohort size and marriage rates. The variation across states confirms the time-series finding that there is a strong and negative relationship between the two variables. In the third step, we provide what is arguably the most convincing evidence on the hypothesis that there is a causal relationship between changes in cohort size and variation in marriage rates. Using an idea proposed by Acemoglu, Autor, and Lyle (2004), we exploit differences in mobilization rates across states during World War II (WWII) as a plausible source of exogenous variation in our variable of interest, cohort size at birth. The results show that an increase in cohort size generates a decline in marriage rates whereas a decline has the opposite effect. The exogenous variation, therefore, gives results that are consistent with the time-series and cross-state variation.

In the second part of the paper, we propose a model that can potentially generate the relationship between changes in cohort size and changes in marriage rates observed in the data. We develop a dynamic search model of the marriage market with the following key feature: women can marry only when young, whereas men can marry when young and old. This modeling choice is based on the fact that men remain fertile longer than women, and on the common insight that an important reason for the existence of marriage is that it is an effective arrangement for the upbringing of children. This key feature of the model implies that men search longer than women for a spouse, which is the pattern that enables us to explain the link between changes in cohort size and changes in marriage rates. Using the model, we prove two results. First, we show that a positive change in cohort size has the effect of reducing the marriage rate and a decline in cohort size has the opposite effect. The model can therefore qualitatively explain the negative relationship between those two variables. We then derive an implication that can be used to test the model. We show that in our dynamic search model, an increase in cohort size has the effect of reducing the age difference between spouses.

In the last part of the paper, we test the ability of the proposed model to explain the observed data. We use the implication derived in the theory part of the paper as our first test. We find that a positive change in cohort size generally reduces the age difference at marriage. The search model is therefore consistent with the data and cannot be rejected. We then estimate the model and evaluate whether it can quantitatively explain the changes in marriage rates observed in the data. We find that the estimated model can explain almost all the variation in marriage rates across cohorts. This result provides additional support in favor of the mechanism we propose as an explanation for the patterns documented in this paper.

The paper proceeds as follows. In Section 2, we discuss related papers. In section 3, we describe the data sets used to derive the empirical results. Section 4 documents our reduced-form findings. In section 5, we develop the search model and derive the two theoretical results. In section 6, we test and estimate the search model. Section 7 concludes.

2 Existing Explanations

In this section, we describe the existing explanations for the variation in the U.S. marriage rates. The discussion emphasizes that the existing explanations for which empirical evidence is provided only apply to particular periods, to a subset of individuals, or are rejected by the data. As a consequence, they cannot account systematically for the historical variation in U.S. marriage rates.

One set of explanations for changes in marriage rates over time focuses on changes in income. A first group of studies emphasize the correspondence between rising incomes during a period of postwar prosperity and the associated marriage boom after World War II. The book by Cherlin (1981) is the main example in this literature. In a related group of studies it has been argued that the decline in income during the Great Depression was the main factor behind the reduction in marriage rates during this period. A positive relationship between income and marriage rates, however, has not been successfully tested over different periods. Hill (2011) rejects the hypothesis that there is a positive correlation between income and marriage rates after 1960. Wolfers (2010) looks at the relationship between marriage rates and recessions for the past 150 years and rejects any pattern between marriage and periods of economic decline. This suggests that an explanation that uses income as the main variable cannot be used as a general theory that can account for the overall variation in marriage rates. In addition, there is a reverse causality problem with this theory that is not addressed in the studies in this literature. It is well documented that married individuals have higher income. It is therefore difficult to determine whether an increase in income causes a rise in the marriage rate or whether an increase in the marriage rate generates higher income levels.

An alternative explanation for the variation in household formation is proposed by Akerlof, Yellen, and Katz (1996). They suggest that one can account for the marriage decline in the seventies using the adoption of new fertility technologies such as the pill and abortion. A potential direct effect could be that the adoption of the new technologies mechanically reduced the number of shut-gun marriages. Akerlof, Yellen, and Katz argue that the adoption of these new methods had also an indirect effect on shotgun marriages. They argue that a decline in the cost of abortion and the increased availability of contraception decreased the incentives of women to obtain a promise of marriage if premarital sexual activity resulted in pregnancy. Those women who are willing to obtain an abortion or who are able to reliably use contraception no longer find it necessary to condition sexual relations on such promises. Those women who want children, who do not want an abortion for moral or religious reasons, or who are unreliable in their use of contraception, may want marriage guarantees but find themselves pressured to participate in premarital sexual relations without any such assurance. As a consequence, the authors argue, marriage rates generally declined. It is immediately evident that this explanation only applies to the seventies when the new fertility technologies were introduced. It cannot account for the large variation in marriage rates observed before and after the seventies.

Greenwood and coauthors have suggested that the decline in the price of appliances explains patterns for several household outcomes and marriage is one of them. In particular, Greenwood and Guner (2009) argue that labor-saving technological progress in the household sector can explain the decline in marriage rates observed during some periods in the past century. Their idea is that technological progress in the household sector makes it easier for singles to maintain their own home, which increases the value of being single and therefore reduces the marriage rate. This theory, however, cannot be used to systematically explain the historical variation in marriage rates. Since technological progress has constantly improved household appliances in the past 100 years, the proposed theory predicts that marriage rates should decline throughout this period. But this is not the case in the data. We will document later in the paper that marriage rates experience large fluctuating during this period, with times in which the fraction married increased sharply and times in which the fraction married declined at a fast rate.

Wilson (1987) proposes a theory of the variation in marriage rates which is based on fluctuations in labor market opportunities. Under the assumption that men are the primary earners, in his book he argues that in periods with limited labor market opportunities there are fewer marriageable men. As a result, the marriage rate declines. Ellwood and Crane (1990) review papers that have tested Wilson's hypothesis and find conflicting evidence. Some papers provide evidence in support of this theory. But there are papers that reject this hypothesis. A couple of examples are Plotnick (1990) and Lerman (1989).

Some papers point at welfare programs and incarceration rates as possible explanations of the changes in the fraction married. In their review, Ellwood and Crane (1990) also evaluate papers that test the link between welfare aid and marriage and they conclude that there is very little empirical support for the proposition that welfare benefits played a major role in the marriage trends in blacks. In particular, they argue that the time-series patterns in welfare benefits are inconsistent with the hypothesis that higher welfare benefits accounted for family changes. Moreover, this explanation applies only to the black population for the majority of the last 100 years since, during those years, the fraction of white families making use of welfare programs was relatively small. Charles and Luoh (2010) study the relationship between incarceration and marriage rates and they find that higher incarceration rates decrease the fraction married in black and white populations. This explanation, however, can only explain the fluctuations in marriage rates after 1980 when large scale incarceration in particular of black men began.

There are two explanations that have some commonality with the one we propose. The first explanation is Easterlin's hypothesis. Easterlin (1987) argues that the relative size of a cohort can explain all the variables that determine the economic and social fortunes of a birth cohort: earnings and unemployment rates, college enrollment rates, divorce, fertility, crime, suicide rates, and marriage. The idea behind this claim is that when income is above the aspiration level for a given cohort, the individuals in a given cohort will be optimistic and therefore will have better economic and social outcomes. If the distribution of income of a cohort is affected by its size, then the size will affect its economic fortunes. Easterlin, however, has provided only indirect evidence in support of his hypothesis and researchers that have attempted to test the general idea behind it have found mixed results. For instance, Pampel and Peters (1995) review papers that have investigated the mechanism outlined by Easterlin and conclude that "the evidence for the Easterlin effect proves mixed at best". More importantly, Easterlin provides no direct evidence on the link between cohort and marriage and on the mechanism behind it, which is the main objective of this paper.

The second explanation that is related to ours is the sex-ratio hypothesis. According to this hypothesis, when the marriage market is characterized by a high sex ratio, measured as the number of available women divided by the number of available men, the marriages rate should decline. Becker (1973) is one of the first attempts to theoretically characterize a possible relationship between the sex-ratio and marriage rates using a matching model. Some papers have tried to test this relationship. For instance, Schoen (1983) considers changes in sex ratios due to two factors: changes in the rate of population growth and differences between women and men in preferences for the age of their spouse. He then evaluates the effect of these changes on marriage rates in the U.S. and finds little effect. In their book, Guttentag and Secord (1983) hypothesize a link between the numerical balance of the sexes and a variety of demographic, social, and psychological consequences among which they incluce marriage rates. To test their hypothesis, Guttentag and Secord use a series of case studies which suggest that declines in the number of women relative to men lead to higher marriage rates. Angrist (2002) looks at variation in immigration rates from different European countries to the U.S. at the beginning of the twentieth century. He exploits the fact that the majority of migrants were men and that most marriages were formed between individuals belonging to the same ethnicity. He finds that ethnicities that experienced lower sex ratios display higher marriage rates. Abramitzky, Delavande, and Vasconcelos (2011) use variation in sex-ratio due to World War I casualties in France. They find that a higher sex ratio is associated with a lower marriage rate for women and a larger marriage rate for men. All these studies look at specific cases and periods. In our paper we show that cohort size explains the observed variation in marriage rates for the entire period we consider. We provide a possible explanation for the relationship between cohort size and marriage rates which relies on a search model. In the model, one of the ways a change in cohort size affects the marriage rate is by changing the number of women relative to men, the sex ratio. The mechanism proposed here is therefore partially related to papers that use changes in the sex ratio to explain changes in marriage rates.[1]

[1]Several papers have studied the relationship between the sex ratio and other economic variables. One example is the paper by Grossbard-Shechtman (1984) where the author analyzes the link between the sex ratio and labor force participation of women. A second example is the paper by Bitler and Schmidt (2011) which shows that

3 Data

For the empirical analysis, we construct two main variables of interest: (1) cohort size, or the number of people born in a particular year, and (2) the share ever married by a given age in each birth cohort.

To measure cohort size, we use the total yearly births recorded in the Vital Statistics of the United States for the years 1909-2011, nationally and by state for white and black populations. Because births before 1960 were only classified as white or non-white, our series for black cohort size includes all non-white births for pre-1960 cohorts. This is not a considerable limitation, as more than 95% of non-white individuals during this time were black. While net migration and early-life mortality affects the final number of marriageable-age adults in each cohort, cohort size provides a very good approximation for this number. Since our analysis begins after the end of large migration waves in the late nineteenth and early twentieth centuries, for most of the period that we consider net migration rates are below 1% of the U.S. population.

The variable share ever married by a certain age in each cohort is constructed using different datasets depending on whether we work with longitudinal or cross-state variation. In the longitudinal analysis, we employ the IPUMS Current Population Survey (CPS) or the Census, depending on the year in which a cohort was born. In the CPS, which covers the period 1962-2011, we observe the age and the marital status of each respondent. We can therefore easily compute the share ever married by age 25, 30, 35, or 40 for each cohort born after a particular year. For instance, for the variable share ever married by age 30, we can use the CPS for all cohorts born on or after 1932; for the variable share ever married by age 40, we can use the CPS for all cohorts born on or after 1922. For cohorts born before those years, we use the 1960, 1970, and 1980 Censuses, which contain information on the marital status and the age at first marriage, a recall variable. Using these two variables, we construct the share ever married by age 25, 30, and 35 for different cohorts by considering all individuals that in a given Census are between the ages of 30 and 45. We use a maximum cutoff age of 45 to avoid potential

there is a causal relationship between the sex ratio and birth rates using differences across states and over time in mobilization rates during the Vietnam war.

measurement errors due to differential mortality rates of married and non-married individuals. For the share ever married by age 40, we use the same procedure with a maximum cutoff age of 50. In the cross-sectional, state-level analysis we use the Census only, as sample sizes in the CPS are too small to provide reliable estimates. In the longitudinal as well as in the cross-state variation, we cannot construct the share ever married for cohort born before 1914 because the 1960 Census is the first one that records the age at first marriage.

Finally, we construct the mobilization rates employed in the instrumental variables regressions using data from the Selective Service System (1956) monographs.

4 Empirical Results

This section is divided into five parts. We first describe the two measures generally used to study the evolution of marriage rates and propose an alternative measure which we believe is better suited to the examination of the marriage rates. We then provide empirical evidence on the relationship between cohort size and marriage rates using longitudinal variation. In the third subsection, we describe our findings obtained using cross-state variation. We then discuss endogeneity issues that may affect the longitudinal and cross-state variation. Finally, we provide evidence that changes in cohort size generate changes in marriage rates using mobilization rates during WWII.

4.1 Different Measures of Changes in the Marriage Rate

When analyzing changes in the marriage rates over time, previous studies have typically employed one of the following two variables: the number of marriages per population; the share of individuals currently or ever married within some age range, e.g. the share of married women between the ages of 18 and 30. As a first contribution, we show that these variables are problematic when used to study the evolution of the marriage market over time.

The first measure, the number of marriages per 1,000 individuals, is described in Figure 1. It is constructed using the U.S. Vital Statistics. The most serious problem with this variable is that it conflates changes in the numerator, new marriages, with changes in the denominator, population. As a result, one draws the incorrect inference about changes in marriage rates whenever the population undergoes any substantive growth or decline due, for instance, to changes in fertility or migration patterns. We will provide evidence that, in the data, there are periods of significant population growth in which this measure reports a decline in marriage rates even if the share ever married by a given age has actually increased. Similarly, we will provide evidence that there exist periods of population decline in which this variable suggests that the economy experienced an increase in marriage rates even though the share ever married by a given age has dropped.

The second measure is presented in Figure 2, where we report the share ever married for a cross-section of women falling in the age range 18 to 30. This variable is constructed using the Census. This measure conflates a different set of effects: changes in the number of people who will ever marry with changes in the age at first marriage. When age at first marriage changes significantly, a researcher will generally draw the wrong inference about variation in overall marriage rates. As we will show, this measure can drop significantly even as the share of married individuals rises.

To avoid the issues that characterize the two variables just described, we propose as an alternative measure, the share of men or women in each cohort ever married by a specific age. This variable has three main advantages. First, it is consistent over time. Second, it does not confound changes in population with changes in the number of marriages. Third, it does not conflates changes in age at first marriage with changes in the number of marriages as long as one chooses the appropriate age in the construction of the variable. For instance, if one is concerned that the share of married women by age 30 in a cohort may be affected by changes in the age at first marriage, it is always possible to construct the same measure by age 35 or 40.

To compare the performance of the three alternative measures, Figure 3 plots the number of marriages per 1,000 people, the share of women ever married between the ages of 18 and 30, and the share of women in each birth cohort married by age 30. Since the first two variables are based on the calendar year whereas our measure is based on the year of birth, to plot them in the same graph we shift forward in time our measure by adding 25 years to the year of birth.

A comparison of the measure that uses the number of marriages per 1,000 individuals with the cohort-based measure makes clear the issue that affects the first variable. Starting from 1946 this variable drops steeply for about fifteen years. Studies that have used this measure have interpreted the drop as a large decline in marriage. The more reliable cohort-based measure, however, tells a different story. During that period, marriage rates increased sharply. An important reason the per-1,000-individuals measure show a large decline is that during those years the U.S. experienced a sharp increase in population with the baby boom. Similarly, during the sixties and the first half of the seventies, the per-1,000-individuals variable displays rapid growth which has been interpreted as a big increase in marriage rates. The cohort measure, however, show that this interpretation is misleading. During this period, marriage rates experienced a slight decline. One important reason the per-1,000-individuals measure displays a rise is because the U.S. population declined during the baby bust. The per-1,000-individuals measure has also problems capturing the actual changes in marriage rates during the nineties when the U.S. population experienced a period of growth. This variable declines, whereas the cohort-based measure shows an increase in marriage rates.

The measure based on the share of individuals married within an age range does a better job in capturing the variation in marriage rates. It follows closely the cohort-based measure for most of the period. It is only during the second half of the eighties and the nineties that the two variables diverge. The measure based on the share of individuals married within an age range would suggest a sustained drop in marriage rates during this period, whereas the cohort-based measure documents a mild increase in the share ever married. The reason for this discrepancy is that during this period the age at first marriage increased, forcing the measure based on the share of individuals married within an age range to display a decline in marriage rates even if they experience a rise.

The evidence reported in this section indicates that the cohort-based measure performs better than the two variables commonly used in the literature. For this reason, in the rest of the paper, we will use it to describe the variation in marriage rates for the past century.

4.2 Change in Marriage Rates Over Time

In this subsection we will provide evidence on the relationship between changes in cohort size and changes in marriage rates using longitudinal variation.

In Figure 4 we plot cohort size and the share never married by age 30 for women and men for all cohorts born between 1914 and 1981. The first panel describes these variables for the white population whereas the second panel plots them for the black population. We plot the share never married, which is computed as one minus the share ever married, to facilitate the comparison between the cohort size and the dynamics in the marriage market. Figure 4 contains two findings that are worth a discussion. First, for all cohorts born before 1960, there is a striking co-movement between a cohort's size and its share never married by age 30. The decrease in the share never married for the small cohorts born in the 1920s and 1930s corresponds to the well-documented marriage boom that starts in the mid-1940s and lasts through the early 1960s. A sharp rise in the share never married begins in 1946 with the first of the post-war baby boom generations. During the first 10 years of the baby boom, total births for whites increased from around two million to nearly three and a half million, with a population increase of similar scale for blacks. Over the same period, share never married by age 30 tripled, until cohort size stabilized. The second finding is that the strong correlation between cohort size and share never married characterizes both the white and the black populations. We emphasize this similarity between the white and black marriage market because in the literature on household formation there is a general perception that the two markets are governed by different rules and therefore should be modeled and estimated using different approaches. Figure 4 suggests, however, that the two marriage markets are similar and respond to changes in cohort size in the same way.

In Figure 4, it is left to explain why the correlation between cohort size and share never married is positive and strong for all cohorts born before 1960, but it weakens for later black and white cohorts. This inconsistency between the early and later cohorts can be resolved by considering the effect of cohabitation. Observe that in Figure 4 cohabiting couples are treated as never-married individuals. Observe also that starting from the eighties, the time when the cohorts born in the sixties became adult, cohabitation starts to be considered a close substitute for marriage. To account for the increase in popularity of cohabitation, in Figure 5 we plot the variables reported in the previous Figures, with the exception that now cohabiting individuals are treated as married individuals. Remarkably, once cohabiting households are accounted for, the relationship between cohort size and household formation resembles again that of the earlier cohorts. Falling cohort sizes in the 1960s and 1970s correspond to a fall in the share never married and not currently cohabiting by 30.

In Figures 4 and 5, we use an age cutoff of 30. The results may therefore be affected by changes in age at first marriage. To address this concern, in Figure 6 we plot cohort size and shares never married and not cohabiting by age 40. Using this new cutoff age, we find the same pattern that is observed in the first two Figures: there is a positive and strong correlation between cohort size and share never married and not cohabiting.

To show these patterns more formally, in Table 1 we record the average response of marriage rates to changes in cohort size at different age cutoffs, ranging from age 25 to age 40. For ease of exposition, for the rest of the paper we will consider the effect of cohort size on the share ever married instead of the share never married. Specifically, in the table each coefficient is the outcome of a separate regression of the log share ever married or currently cohabiting on log cohort size. There are three results that are worth discussing. First, elasticities recorded in Table 1 are highest at 25, and gradually decrease with age, for both sexes and both races. This finding suggests that changes in cohort size are associated with two effects: (i) a change in the eventual share ever married or cohabiting; (ii) a change in the age at first marriage with an increase in cohort size being associated with a higher age at first marriage. This finding also indicates that the coefficient that is better able to isolate the relationship between variation in cohort size and variation in marriage rates from changes in age at first marriage is the one that uses 40 as a cutoff age. The second result is that the effect of cohort size is quantitatively large and is highest for black women. Across all groups, the share ever married or cohabiting by 40 decreases by about 0.66% to 4.53% for every 10% increase in cohort size. In percentage points, this amounts to an approximately 0.6 to 3.8 point decrease in the share of individuals ever married or cohabiting by 40, a large effect. The last finding we wish to emphasize is that the cohort size variable explains a large fraction of the variation observed in marriage rates. For instance, when we use 40 as the cutoff age, the R-squared is between 0.58 and 0.85. [2]

To summarize, our results indicate that in the time-series data there is a strong and negative relationship between marriage rates and cohort size. The results also indicate that changes in cohort size account for a large fraction of the time series variation in marriage rates for both black and white populations. In the following subsections, we further explore the empirical link between cohort size and the marriage rate. Because cohabitation has become a close substitute for marriage since the early eighties, for the rest of the paper we will continue to use the same adjusted measure of household formation: the share ever married or cohabiting by a given age. Unless we specifically note otherwise, when we use the shorthand "ever married" we refer to those ever married or currently cohabiting.

4.3 Change in Marriage Rates Across States

In this subsection we provide additional evidence on the relationship between cohort size and household formation rates by using variation across states. The idea is that if changes in cohort size generate variation in marriage rates, we should observe such an effect not just across time but also across geography. We should observe that states with large increases in cohort sizes experience large drops in marriage rates relative to other states and vice versa.

To use cross-state variation, there is one issue we need to address that was not present in the longitudinal analysis. Changes in the size of a cohort at the time individuals are of marriageable-age is endogenous because it is partially driven by migration decision. These decisions are generally related to differences across states in economic and social conditions which also affect marriage rates. To exacerbate the problem, migration can be sex-biased. It can therefore skew sex ratios and affect marriage rates directly. This concern is minor in the longitudinal variation because there is little migration from abroad for white and black during

[2] Note that we are working with non-stationary time series, and must therefore verify that the series are cointegrated to eliminate worries about spurious regression. A Johansen cointegration test rejects the null hypothesis that the series are not cointegrated at the one-percent level. Therefore, our OLS results are consistent and estimate a meaningful (non-spurious) relationship.

this period. To avoid these potential sources of endogeneity, we perform the cross-state analysis by using the total births in a given year and state. This variable has two useful features. First, it approximates the number of adults from that cohort that will be in the marriage market of that state twenty or thirty years later. Second, it is not affected by the endogeneity concern discussed above.

The two main variables needed for the analysis, cohort size at birth and share ever married, are constructed as follows. For cohort size at birth, we use births by state and race from the Vital Statistics of the United States. We start with the 1940 cohort because it is the first decennial cohort for which we have total births by state for whites and non-whites without missing data for a substantial number of states. For the share ever married, we use the 1970, 1980, 1990, and 2000 Censuses. We start with the 1970 Census because it is the one in which we observe the 1940 cohort at age 30. We end the analysis with the 2000 Census because it is the last one that is publicly available. In the 1990 and 2000 Censuses, the recall variable age at first marriage is not available. As a consequence, we can construct the share ever married by age 30 only for cohorts that were 30 at the time of the Census, namely for the cohorts born in 1940, 1950, 1960, and 1970. Similarly, we can compute the share ever married by age 40 only for the cohorts born in 1940, 1950, and 1960. For each one of these cohorts, we construct the variable share ever married by first assigning each individual in the Census to a particular state based on the birthplace. We then compute the share of individuals born in that state who was ever married.

Using the two variables, for each cohort and state, we construct ten-year differences for the log of share ever married and the log of cohort size. Because of the small sample size, we then pool all cross-sections and regress the differences in log share ever married on the differences in log cohort size. To control for general time trends, we add year fixed effects to the regression. Table 2 presents the results of the cross-state regressions separately by race and gender. For the share ever married by 30, the size of the coefficients is between -0.064 and -.092 and strongly significant. The results obtained using the share ever married by 40 are similar, with coefficients between -0.037 and -0.077. Two results are worth emphasizing. First, similarly to our findings obtained with the time series evidence, the coefficients are higher for black men and women and decline with age. Second, the effects measured using the cross-state variation are smaller than those measured in the longitudinal regressions. One potential reason for this finding is that, because of inter-state migration, state-level births for a given cohort are a noisy measure of the number of marriageable-age adults in that cohort and state. The coefficients on cohort may therefore be biased downward. Additionally, if changes in cohort size have cumulative effects over time, as we will argue in the model section, regressions that use ten-year differences may not fully capture these effects.

4.4 Potential Endogeneity Concerns

The findings obtained using cross-state regressions strongly corroborate the negative relationship between changes in cohort size and changes in marriage rates observed when longitudinal variation was used. However, there are reasons that prevent a causal interpretation of the relationship. While using number of births as our independent variable allows us to reasonably avoid reverse causality problems as well as important endogeneity concerns due to migration, one may nevertheless worry about omitted variables: state-level characteristics that drive changes in birth rates for a particular cohort as well as changes in marriage decisions of individuals that belong to that cohort 20 to 30 years later. Such omitted variables would have to be highly persistent shocks that affect growth in births in a given year as well as growth in subsequent marriage rates about 20 to 30 years later.

It is not easy to think of variables that fit this description. Potential examples include highly persistent productivity shocks which cause wages to grow more rapidly in some states over time, affecting both birth rates at the time of the initial shock and marriage rates two or three decades later. A positive trend in men's earnings in some states fits this description. If children are a normal good, states with such positive trends may see both increased births in 1950 relative to 1940 as well as a greater number of marriageable men in 1980 relative to 1970. Alternatively, improving fertility technologies may have had a differential effect in states that are strongly religious or have stronger preferences for forming a family compared to states that do not. In states with weaker preferences for family, one might expect depressed birth rates as well as lower marriage rates in the future.

Note that in these and most credible cases we would typically expect an increase in both births and subsequent marriage rates or a decline in both variables. This bias would work against our favor and would result in a positive coefficient on cohort size, which is not what we find. Nevertheless, without exogenous variation in cohort size we cannot entirely eliminate the possibility that some biases could work in our favor.

4.5 Instrumental Variables Strategy

To isolate a plausibly exogenous source of variation in births, we employ an idea proposed originally by Acemoglu, Autor, and Lyle (2004) and exploit differences in mobilization rates across states during World War II. From 1940 to 1947, over 10 million men were inducted into the military, with the largest bulk of American soldiers deployed starting in 1943 and 1944, most for the duration of the war. More importantly, the mobilization rates differed substantially across states during the war, ranging from approximately 41.2 percent to 54.5 percent of young men ages 18-44. The induction rates, the fraction of individuals that were drafted, display a similarly high degree of cross-state variation, ranging from approximately 26.4 percent to 35.9 percent. We will show below that this cross-state variation introduced substantial differences in birth rates across states.

4.5.1 Data Description and the Effect of Mobilization on Births

In the instrumental variable regressions, we focus exclusively on white individuals for two reasons. First, black populations in the 1940s were not sufficiently large in many states outside the South to construct reliable marriage rates by cohort using Census samples. Second, mobilization rates for black men were very low as discussed in Acemoglu, Autor, and Lyle (2004). For these reasons, the variables needed in the analysis are constructed using only whites. We construct the variable cohort size at birth for each state by using the National Vital Statistics for white individuals only from 1937, five years before the large-scale mobilization, to 1950, five years after the end of the war.

Our mobilization rate measure is also computed for whites only using the Selective Service's Special Monograph (1956) for men between the ages of 18 and 44. The Selective Service's Special Monograph provides breakdowns by race of draft registrations and inductions, but not of enlistments. We therefore construct the mobilization rate measure as the number of white men ages 18-44 inducted into the armed forces divided by the number of white men ages 18-44 registered. The latter includes virtually all men ages 18-44 between 1940 and 1945. A benefit of constructing the mobilization measure in this particular way is that the key source of variation in participation in the armed forces stems from randomness generated by the cross-state differences in induction processes of draft boards, rather than by differences across states in men's voluntary enlistment rates. We will discuss later the factors that generated heterogeneity in the induction processes.

To construct the share ever married for each state and for each cohort born between 1937 and 1950, we can only use the 1970 and 1980 Censuses. We cannot use later Censuses because to compute the marriage variable for all those cohorts we need the recall variable age at first marriage, which is not available after 1980. We cannot use the CPS because its sample size is too small to construct the marriage measure at the state level. Since we are limited to the 1970 and 1980 Censuses, we can only perform the instrumental variable analysis for the share ever married at age 30. For the share ever married at age 40, we do not observe in those Censuses cohorts born after 1940, since individuals that are 40 or older in the 1980 Censuses were born in 1940 or earlier. However, this should not be an important limitation considering that we have shown, using longitudinal and cross-state variation, that the effect of cohort size on marriage rates is similar using the share ever married at age 30 and at age 40.

We use the variables just described to illustrate in Figures 7, 8, and 9 that the effect of higher mobilization rates on births can be divided into three parts. First, during the first period of mobilization, fathers were granted a delay in deployment. As a consequence, during the first phase of WWII, we would expect higher birth rates for men with higher probabilities of being deployed and consequently in states with a larger fraction of them. Figure 7 describes this “anticipation effect” for the years 1940 and 1942. As predicted, there is a positive correlation between the mobilization rates and changes in log births, with a regression coefficient of about 0.8. To show that the anticipation effect is not a cross-state difference that existed before the war and is therefore independent of mobilization, in Figure 10 we report the same correlation for the years 1935 to 1940. Our results indicate that before the war the correlation between mobilization rates and changes in log births, while insignificant, has the opposite sign of what the anticipation effect predicts. Second, during the second phase of the war when the majority of U.S. men were deployed, in states with higher rates of mobilization, the absence of prime-age men should be associated with a larger drop in births. We will refer to this second effect as the “incapacitation effect”. Figure 8 shows the correlation between mobilization rates and changes in log births from 1943 to 1945. As expected, the correlation is negative. In states with high mobilization, the greater absence rate of potential partners leads to a greater drop in births. The coefficient on the regression line suggests that a one percentage point increase in the mobilization rate is associated with 1.26 percent lower cohort growth. Finally, we would expect the end of the war to affect the birth rates. We will refer to this effect as the “return effect”. There are two opposing forces that can generates the return effect. First, we would expect a larger increase in birth rates in states with high mobilization rates when the deployed men who survived the war returned home and started to make up for the missing years. Second, many of the deployed men died during the war. As a consequence, the birth rate may have declined faster in states with a larger share of deployed men because of the lower number of prime age male partners. Figure 9 plots the correlation between mobilization rates and changes in log births for 1945 and 1946. We find that the post-war recovery in births in states with high mobilization rates is greater than the recovery in low-mobilization states. A one percentage point increase in mobilization rates is associated with a 0.58 percent increase in the growth in births. This suggests that the first component of the return effect dominates.

4.5.2 Specification

We use the cross-state variation in mobilization rates jointly with a standard instrumental variable estimator to measure the causal effect of changes in cohort size on marriage rates. In the first stage, we instrument log cohort size using the state's mobilization rates by estimating the following equation:

$$logcohortsize_{c,s} = \sum_s \pi_s + \sum_c \gamma_c + \sum_c \gamma_c \cdot mobilizationrate_s + X_{c,s}\beta + \nu_{c,s}, \tag{1}$$

where π_s is a set of state fixed effects, γ_c is a set of birthyear fixed effects, $X_{c,s}$ is a set of 1940 state demographic characteristics interacted with time that allows for potentially different time trends in states with different baseline characteristics. At the end of this subsection we will discuss in details which variables enter $X_{s,c}$.

We consider four different specifications of equation (1) depending on how the mobilization rates and the birthyear fixed effects enter the first stage. In the baseline specifications, we include the interactions between the overall mobilization rate in a state and year dummies from 1941 to 1947, year fixed effects, and time trends interacted with baseline state characteristics. In the second specification, we add two additional sources of variation in mobilization rates to potentially generate more precise estimates of the anticipation, incapacitation, and return effects: the cumulative year-by-year mobilization rates at the state level from 1941 to 1944 and the war casualty rate by state. Observe that the anticipation effect is based on the prediction during the first part of WWII about the fraction of men that would have been deployed and not on the actual share. It is therefore important that in the second specification we include both the overall mobilization rate interacted with year dummies, which better capture the prediction of future deployments, and the cumulative mobilization rates. The third specification adds to the first one birthyear fixed effects that vary by region. The inclusion of these variables allows us to control for the possibility that part of the cross-state variation in mobilization rates is generated by underlying differences in time trends across regions. In the last specification, we include both region-birthyear fixed effects and cumulative year-by-year mobilization rates.

In the second stage, we estimate a model similar to the one we introduced in the previous subsections:

$$y_{c,s} = \sum_s \pi_s + \sum_c \gamma_c + \phi \cdot logcohortsize_{c,s} + X_{c,s}\beta + \epsilon_{c,s} \qquad (2)$$

where $y_{c,s}$ is the share of individuals ever married who were born in year c and state s and ϕ is our coefficient of interest.

To determine whether the cross-state variation in mobilization is arguably exogenous, it is important to understand why different states experienced dissimilar induction rates. There are five main reasons behind the heterogeneity in mobilization rates. First, the registration of all men 18 to 60 as well as their classification, deferrals, and eventual induction into the armed forces was overseen by the the civilian-led Selective Service System, which was created during peacetime in September 1940. The Selective Service was highly decentralized and administered by more than 6,000 local draft and appeal boards throughout the nation. The local boards were staffed by a small number of volunteers who received directives from state-level and national offices, but had significant autonomy in the classification process and in granting deferrals. As a consequence, the first source of variation is generated by the idiosyncratic differences in the behavior of local draft boards. The second source of variation is created by differences in age structure across states as states with a larger fraction of older men had lower rates of inductions. These two sources of variation are the type of variation we would like to have in our analysis because they are unlikely to be correlated with marriage patterns 20-30 years later. A third source of heterogeneity is produced by differences in ethnic compositions across states. States with high concentration of Germans, Japanese, and Italians had lower mobilization rates since members of these ethnic groups typically were not allowed to serve. Observe that if one of these ethnicities has stronger preferences for large families, states in which this ethnicity is concentrated will have higher birth rates today and higher marriage rates in 20-30 years. As a consequence, a potential bias will work against our favor if we find evidence of a negative causal relationship between cohort size and marriage rates. The fourth source of variation is generated by differences in occupational structures as workers employed in industries central to the war effort were more likely to be exempted. Finally, exemptions for farmers played a role in the induction process because food production was a priority for the U.S. at this time. These last two sources of variation could pose a potential problem for the empirical strategy if they are associated with systematic differences in income across states.

To evaluate whether this is the case, we first group states into three groups based on armed forces induction rates: low-mobilization states, with induction rates below 28.8 percent; medium-mobilization states, with induction rates between 28.8 and 31.2 percent; and high-mobilization states, with induction rates greater than 31.2 percent. We then compare the mean for the main variables of interest. Notice that we only employ inductions in the construction of our mobilization rates. Our rates are, therefore, lower than those constructed by Acemoglu, Autor, and Lyle (2004). The results are presented in Table 3. They indicate that of all the 1940 measures we consider only two variables are statistically different across groups. The first variable that is different across groups is mean age, which is lower in high-mobilization states since younger men were more likely to get drafted. Second, the difference in mean years of education between high- and low-mobilization states is statistically significant, with low-induction rate states recording about 0.75 more years of education. The share of men who are farmers is higher in low-mobilization states, as one would expect, but the difference is not statistically significant at the 10%-level. There is virtually no difference across groups in average log income.[3]

In all our specifications, we control for state fixed effects. As a consequence, if the educational differences between high- and low-mobilization states are constant over time, they will not affect our estimates. However, it is possible that states may also have differing trends in this independent variable. To address this concern, we include in the set of control variables $X_{s,c}$ average years of education in 1940 interacted with a fourth-order polynomial in time. To be on the safe side, we include in $X_{s,c}$ also the share of farmers and log income in 1940 interacted with the polynomial in time.

[3] Our summary statistics differ somewhat from those reported in Acemoglu, Autor, and Lyle (2004) for two reasons. First, we use a different measure of the mobilization rate (inductions only). Second, our sample includes white individuals only. Acemoglu et al. do not separate their sample by race. We find similar results when we use the full sample and mobilization rates from Acemoglu et al. (2004). These results are available online at http://www.econ.ucla.edu/mazzocco/research.htm.

4.5.3 Results

Table 4 reports the outcome of the first-stage regression. In column (1), we describe the results for the baseline specification. A positive coefficient on the mobilization rate interacted with the year 1942 is in line with the idea that men attempted to delay their deployment by becoming fathers, but it is not significant at conventional levels with a p-value equal to 0.16. The coefficient estimate on the mobilization rate becomes zero for 1943 and turns negative and statistically significant at the five-percent level for 1944. As expected, the negative coefficient grows larger in magnitude in 1945 and becomes significant at the one-percent level. Some of the effect persists into 1946, which is in line with the fact that soldiers remained in active service for up to 6 months after the end of the war. In this specification, we therefore observe a strongly significant and negative incapacitation effect from 1944 to 1946. The coefficient becomes statistically insignificant thereafter, implying that we do not observe a net return effect. In column (2), we present the results for the specification that includes the cumulative mobilization rates. The only effect of introducing this additional variation is that the year-by-year mobilization rate absorbs most of the effect for 1944.

Column (3) contains the estimation results for the first specification that controls for region fixed effects. In this case, the anticipation effect becomes statistically significant, with a coefficient for 1942 that is positive and large. The incapacitation effect, however, disappears with the coefficients for 1944 and 1945 now statistically equal to zero. In column (4), we introduce the additional year-by-year variation in the mobilization rate and the variation in the casualty rate. In this specification, the anticipation effect, the incapacitation effect, as well as the return effect appear to be important. The anticipation effect is captured by a positive and statistically significant coefficient for 1942. The incapacitation effect is evident in a negative and statistically significant coefficient on the cumulative mobilization rate for 1944. Finally, the return effect is captured by a large and negative coefficient on the casualty rate.

We can now discuss the results of the second stage regression. Table 5 reports the estimated coefficients obtained by regressing log share ever married by 30 on log cohort size separately for men and women and for all four specifications. The effect we measure is negative and strongly significant. We find that a one percent increase in cohort size generates a 0.027% to 0.053% reduction in the share of women ever married and a 0.023% to 0.029% decrease in the share of men ever married by 30. As in the previous sections, we find a larger effect for men than for women, although the coefficients overall are smaller than the ones recorded in the cross-sectional regressions. This is not surprising. Even though the reduction in births due to mobilization measured in our first stage regression is potentially large, the period during which cohort size was affected during the war is very short, potentially only three or four years. Changes in cohort size over a longer time horizon will generally have stronger effects that cumulate over time. The difference between the short and long run effects should account for the larger effects measured in the previous section. Nevertheless, we believe that the fact that such a short-term reduction in cohort size has a noticeable and strongly significant effect provides compelling evidence for the causal effect of changes in cohort size on changes in marriage patterns.

5 A Dynamic Search Model of the Marriage Market

In this section we develop a simple dynamic search model of the marriage market. We then show that the model can match the negative casual relationship between cohort size and marriage rates observed in the data. Finally, we derive an implication that will be used in the next section to test the model. A possible alternative to the search model we consider in this section is a matching model of the type employed by Gale and Shapley (1962), Becker (1973), Becker (1974), Mortensen (1988), Bergstrom and Bagnoli (1993), Peters and Siow (2002), Choo and Siow (2006), Chiappori, Iyigun, and Weiss (2009), Hitsch (2010), and Iyigun and Walsh (2007). We decided to work with a search model because it allows us to better capture the dynamic nature of the marriage market which is essential to explain the link between cohort size and marriage rates.

The model characterizes an economy populated by $T+1$ overlapping generations of men and women. In each period $t = \{0, ..., T\}$, a new generation is born and lives for $T+1$ periods. Men and women can be either single or married. If an individual is married she or he makes no choice. If in period t an individual of gender i and age a is single, she or he meets a potential spouse with probability $\theta_{a,t}^{i}$. The two spouses then decide whether to marry with the objective of maximizing their lifetime utility.

We now introduce the main assumption of the model. We assume that women meet a man with a positive probability only in their first period of life, while men meet a potential spouse with a positive probability in their first two periods of life. This assumption is justified by the following two facts. First, women are fertile only in their first part of their adult life, whereas men are fertile for most of their adult life. Second, there is a common belief that one of the main reasons for marriage is to give birth to and nurture children. These two facts imply that the value of marriage for a woman declines faster than the value for a man. Our assumption is a special case of an economy in which the value of marriage for women and men follow this pattern. Our main assumption has two implications. First, the marriage market is populated by women of age 0 and by men of age 0 and 1. Second, women cannot change their marital status after the first period and men cannot change it after the second period. Allowing women to marry for more than one period with a declining value of marriage and men to marry for more than two periods makes the model more complicated without changing the qualitative nature of the results.

The within-period utility of being single will be denoted by δ, whereas the within-period utility of being married for the couple as a whole will be denoted by η. The value of being married is drawn from a distribution $F(\eta)$ which does not depend on the age of the couple or on time. The utility from future periods is discounted at the discount factor $\beta \leq 1$. We will assume that the value of being single is constant across individuals and over time. As a consequence, if an older man chooses to be single in the second period, his lifetime utility takes the following form:

$$v_{1,t}^{m} = \sum_{t=0}^{T-1} \beta^{t} \delta = \frac{1-\beta^{T}}{1-\beta} \delta.$$

Similarly, if a woman decides to stay single in her first period of life, her lifetime welfare can be computed as follows:

$$v_{0,t}^{w} = \sum_{t=0}^{T} \beta^{t} \delta = \frac{1 - \beta^{T+1}}{1 - \beta} \delta.$$

If two potential partners decide to marry, the within-period utility they have drawn is also the utility they will experience in each period for the rest of their life. The lifetime utility of a couple of individuals who are both of age 0 and have drawn a value η in period t can therefore be written as follows:

$$v_{0,0,t} = \sum_{\tau=0}^{T} \beta^{\tau} \eta = \frac{1 - \beta^{T+1}}{1 - \beta} \eta.$$

If the couple is composed of on older man and a woman, the man will die one period earlier. As a consequence, their lifetime utility takes the following form:

$$v_{0,1,t} = \sum_{\tau=0}^{T-1} \beta^{\tau} \eta + \beta^{T} \delta = \frac{1 - \beta^{T}}{1 - \beta} \eta + \beta^{T} \delta.$$

We will assume that the couple can freely divide the gains from marriage and that its lifetime utility is split between the two spouses using a Nash bargaining solution. For a couple composed of a woman of age 0 and a man of age 1, the share received by the man in period t is

$$w_{1,t}^{m}(\eta) = v_{1,t}^{m} + \gamma \left[v_{0,1,t} - v_{1,t}^{m} - v_{0,t}^{w} \right] = v_{1,t}^{m} + \gamma_{m} \left[\frac{1 - \beta^{T}}{1 - \beta} \eta + \beta^{T} \delta - v_{1,t}^{m} - v_{0,t}^{w} \right], \qquad (3)$$

where the parameter $\gamma \in [0, 1]$ captures the bargaining power of the man and $v_{1,t}^{m}$ and $v_{0,t}^{w}$ are the value of being single in this and future periods that were computed above. A similar equation can be derived for the woman. To make it harder for our model to explain the variation observed in the data, we will assume that γ is independent of market conditions.

We can now solve the model starting with the decisions of a man of age 1 in period t. With probability $\theta_{1,t}^{m}$, he meets a woman and they marry if their joint lifetime utility from being married is greater than the sum of their lifetime utilities if they choose to stay single. To determine the match quality η above which the couple will choose to marry, observe that if a man of age 1 and a woman of age 0 decide to remain single, they will be single for the rest of their life. As a consequence, they will marry if and only if

$$\eta\frac{1-\beta^{T}}{1-\beta}+\delta\beta^{T}\geq\delta\frac{1-\beta^{T}}{1-\beta}+\delta\frac{1-\beta^{T+1}}{1-\beta}=2\delta\frac{1-\beta^{T}}{1-\beta}+\delta\beta^{T}.$$

This implies that the reservation value for marriage between a woman and a man of age 1 is

$$\underline{\eta}_{1,t}=2\delta.$$

We can now derive the expected value function for an older man before he enters the marriage market. If in period t this man meets a woman and draws a value η, Nash-bargaining implies that he receives the following share of the couple's lifetime utility:

$$w_{1,t}^{m}(\eta)=\delta\frac{1-\beta^{T}}{1-\beta}+\gamma\left[\eta\frac{1-\beta^{T}}{1-\beta}+\delta\beta^{T}-2\delta\frac{1-\beta^{T}}{1-\beta}-\delta\beta^{T}\right]=\left[\delta+\gamma\left(\eta-2\delta\right)\right]\frac{1-\beta^{T}}{1-\beta}.$$

As a consequence, the expected value function of an older man can be written in the following form:

$$v_{1,t}^{m}=E\left[\delta+\gamma\left(\eta-2\delta\right)\left|\eta\geq\underline{\eta}_{1}\right.\right]\frac{1-\beta^{T}}{1-\beta}\left(1-F\left(\underline{\eta}_{1}\right)\right)\theta_{1,t}^{m}+\delta\frac{1-\beta^{T}}{1-\beta}F\left(\underline{\eta}_{1}\right)\theta_{1,t}^{m}+\delta\frac{1-\beta^{T}}{1-\beta}\left(1-\theta_{1,t}^{m}\right)$$

It is composed of three parts. The first part describes the value for the older man of meeting a woman with a match quality η sufficiently high that the couple will choose to marry multiplied by the corresponding probability. The second part characterizes the value of meeting a woman with a match quality η that is below the reservation value $\underline{\eta}_{1}$ times the probability of this event. Finally, the last part captures the value of not meeting a woman in the current period multiplied by the probability. By replacing $\underline{\eta}_{1}=2\delta$, by dividing both sides of the equation by $\frac{1-\beta^{T}}{1-\beta}$, and by simplifying some of the terms, we obtain the following equation for the value function:

$$v_{1,t}^{m}\frac{1-\beta}{1-\beta^{T}}=\gamma\left\{E\left[\eta\left|\eta\geq2\delta\right.\right]-2\delta\right\}\left(1-F\left(2\delta\right)\right)\theta_{1,t}^{m}+\delta. \tag{4}$$

We are now in the position to consider the decision of a young man. This individual meets a potential spouse with probability $\theta_{0,t}^{m}$ and they marry if their joint lifetime utility is greater than the sum of their lifetime utilities if they choose to be single in this period, i.e. if

$$\eta\frac{1-\beta^{T+1}}{1-\beta}\geq 2\delta+\beta v_{1,t}^{m}+\beta\delta\frac{1-\beta^{T}}{1-\beta},$$

where the first term on the right hand side is the joint value of being single in this period, the second term is the man's discounted expected value function for next period if he chooses to stay single today, and the third term is the woman's discounted value from next period onward if she chooses to stay single today. The reservation value for a man of age 0 can therefore be written as follows:

$$\underline{\eta}_{0,t}=2\delta\frac{1-\beta}{1-\beta^{T+1}}+\beta v_{1,t}^{m}\frac{1-\beta^{T}}{1-\beta^{T+1}}+\beta\delta\frac{1-\beta^{T}}{1-\beta^{T+1}}$$

We can now substitute for the expected value function of an older man using equation (4) and simplify some of the terms to obtain the following equation for the reservation value of a young man:

$$\underline{\eta}_{0,t}=2\delta+\beta\frac{1-\beta^{T}}{1-\beta^{T+1}}\gamma\left\{E\left[\eta\left|\eta\geq 2\delta\right.\right]-2\delta\right\}\left(1-F\left(2\delta\right)\right)\theta_{1,t}^{m}. \tag{5}$$

Using $\underline{\eta}_{0,t}$, one can derive the expected value for a woman and men. They are presented in Appendix A.4.

5.1 Steady State

In this subsection, we use the reservation values discussed above to solve for the steady state equilibrium in the marriage market. We will use it to derive a couple of theoretical results and as an input for the structural estimation of the model which will be discussed in the next section.

To solve for the steady state equilibrium, we have to derive the probability that a younger man meets a woman $\theta_{0,t}^{m}$ and the corresponding probability for an older man $\theta_{1,t}^{m}$. Let N_{t}^{i} be the number of individuals of gender i and age t who are present in the marriage market. Then $\theta_{0,t}^{m}$

and $\theta_{1,t}^m$ can be derived by noting that

$$\theta_{0,t}^m = \theta_{1,t}^m = \frac{N_{0,t}^w}{N_{0,t}^m + N_{1,t}^m}. \tag{6}$$

The number of individuals of age 0 is exogenously given by the cohort size of a generation. However, the number of older men in the marriage market $N_{1,t}^m$ is endogenously determined by the decisions of younger men. As a consequence, to derive $\theta_{0,t}^m$ and $\theta_{1,t}^m$ we need to solve for $N_{1,t}^m$. This variable can be computed as the number of younger men who did not meet a woman at $t-1$ plus the number of younger men who met a woman at $t-1$ but draw a match quality η lower than the reservation value, i.e.

$$N_{1,t}^m = N_{0,t-1}^m \left(1 - \theta_{0,t-1}^m\right) + N_{0,t-1}^m \theta_{0,t-1}^m F\left(\underline{\eta}_{0,t-1}\right) = N_{0,t-1}^m \left(1 - \theta_{0,t-1}^m \left(1 - F\left(\underline{\eta}_{0,t-1}\right)\right)\right). \tag{7}$$

We can now replace for $\theta_{0,t-1}^m$ using (6) and obtain the following equation for $N_{1,t}^m$:

$$\begin{aligned} N_{1,t}^m &= N_{0,t-1}^m \left(1 - \frac{N_{0,t-1}^w}{N_{0,t-1}^m + N_{1,t-1}^m}\left(1 - F\left(\underline{\eta}_{0,t-1}\right)\right)\right) \\ &= N_{0,t-1}^m \left(\frac{N_{0,t-1}^m + N_{1,t-1}^m - N_{0,t-1}^w \left(1 - F\left(\underline{\eta}_{0,t-1}\right)\right)}{N_{0,t-1}^m + N_{1,t-1}^m}\right). \end{aligned}$$

In a steady state equilibrium, the cohort size $N_{0,t}^w$ and $N_{0,t}^m$ and the number of older men in the marriage market $N_{1,t}^m$ are constant over time. We therefore have that

$$N_1^m = N_0^m \left(\frac{N_0^m + N_1^m - N_0^w \left(1 - F\left(\underline{\eta}_{0,t-1}\right)\right)}{N_0^m + N_1^m}\right).$$

We can now solve for N_1^m and obtain

$$N_1^m = \sqrt{\left(N_0^m\right)^2 - N_0^m N_0^w + N_0^m N_0^w F\left(\underline{\eta}_0\right)}.$$

Observe that generally men and women have identical cohort size, i.e. $N_{1,t}^m = N_{1,t}^w = N_{1,t}$.[4] In this case the solution for N_1^m simplifies to

$$N_1^m = N_0 F\left(\underline{\eta}_0\right)^{\frac{1}{2}}.$$

If we substitute N_1^m back into θ_j^m, we have

$$\theta_0^m = \theta_1^m = \frac{N_0^w}{N_0^m + \sqrt{\left(N_0^m\right)^2 - N_0^m N_0^w + N_0^m N_0^w F\left(\underline{\eta}_0\right)}}.$$

If men and women have identical cohort size, θ_j^m simplifies to

$$\theta_0^m = \theta_1^m = \frac{N_0}{N_0 + N_0 F\left(\underline{\eta}_0\right)^{\frac{1}{2}}} = \frac{1}{1 + F\left(\underline{\eta}_0\right)^{\frac{1}{2}}}.$$

To determine the reservation value of younger men in steady state, we can substitute for θ_1^m in the equation that determines the reservation value (5). We can then derive, for the case in which $N_0^m \neq N_0^w$, the following equation for the steady state reservation value:

$$\underline{\eta}_{ss} = 2\delta + \beta \frac{1-\beta^T}{1-\beta^{T+1}} \gamma \left\{E\left[\eta \left| \eta \geq 2\delta\right.\right] - 2\delta\right\} \left(1 - F\left(2\delta\right)\right) \frac{N_0^w}{N_0^m + \sqrt{\left(N_0^m\right)^2 - N_0^m N_0^w + N_0^m N_0^w F\left(\underline{\eta}_{ss}\right)}},$$

If $N_0^m = N_0^w$, the equation simplifies as follows:

$$\underline{\eta}_{ss} = 2\delta + \beta \frac{1-\beta^T}{1-\beta^{T+1}} \gamma \left\{E\left[\eta \left| \eta \geq 2\delta\right.\right] - 2\delta\right\} \left(1 - F\left(2\delta\right)\right) \frac{1}{1 + F\left(\underline{\eta}_{ss}\right)^{\frac{1}{2}}}.$$

Observe that $F\left(\underline{\eta}\right)$ is monotonically increasing in $\underline{\eta}$. As a consequence, there is a unique solution for $\underline{\eta}_{ss}$. Moreover, if men and women have identical cohort sizes, the steady state reservation value is independent of N_1^m and N_1^w. The following Proposition summarizes the result.

[4] This is not the case if men or women are more likely not to be in the marriage market for particular reasons. For instance, African-American men are more likely than African-American women to be in jail during their marriage years. As a consequence, the relevant cohort size for African-American men is smaller than the corresponding cohort size for women.

Proposition 1 *In steady state, there is a unique reservation value for marriage $\underline{\eta}_{ss}$. It does not depend on cohort size if $N_0^m = N_0^w$.*

5.2 An Unexpected Shock to Cohort size

We will now consider the effect of a shock to cohort size on the fraction of individuals that choose to marry. We will focus on the case in which the shock is unexpected. Similar results apply if the shock is known with certainty. We will show two results. The first result is that a positive shock to cohort size reduces the fraction of women in a given cohort who choose to marry. The second result provides a testable implication for the model and it establishes that an increase in cohort size reduces the average age difference between spouses. In the remaining part of the section we will consider the case in which $N_0^m = N_0^w = N_0$. As a consequence, the results do not apply to African-Americans since for this population the jail rate and mortality rate are higher for men of marriage age than for women.

Suppose the economy is in steady state when it is hit by an unexpected shock in period $t = \tau$ that changes permanently the women's cohort size from N_1^w to $N_1^w + \Delta$ and the men's cohort size from N_0^m to $N_0^m + \Delta$. We consider the case of a permanent shock because in the data changes in cohort size tend to be persistent and even reinforcing. According to equation (6), the probabilities $\theta_{j,t}^m$ take the following form:

$$\theta_{0,t}^m = \theta_{1,t}^m = \frac{N_{0,t}}{N_{0,t} + N_{1,t}^m} \qquad \text{if } t < \tau$$

and

$$\theta_{0,t}^m = \theta_{1,t}^m = \frac{N_{0,t} + \Delta}{N_{0,t} + \Delta + N_{1,t}^m} \qquad \text{if } t \geq \tau.$$

Consider the period in which the shock is realized and notice that $N_{1,\tau}^m$ are the men born in period $\tau - 1$ who did not marry when young. As a consequence, $N_{1,\tau}^m$ equals the number of older men in steady state, i.e. $N_{1,\tau}^m = N_{0,\tau-1} F\left(\underline{\eta}_{ss}\right)^{\frac{1}{2}} = N_0 F\left(\underline{\eta}_{ss}\right)^{\frac{1}{2}}$. Substituting for $N_{1,\tau}^m$ in the probabilities $\theta^m_{j,t}$, we have that in period τ

$$\theta^m_{0,\tau} = \theta^m_{1,\tau} = \frac{N_0 + \Delta}{N_0 + \Delta + N_0 F\left(\underline{\eta}_{ss}\right)^{\frac{1}{2}}} = \frac{1}{1 + \frac{N_0}{N_0+\Delta} F\left(\underline{\eta}_{ss}\right)^{\frac{1}{2}}}.$$

The previous equation implies that a positive cohort shock Δ increases the probability that a man of any age meets a woman, whereas a negative cohort shock has the opposite effect. Observe that in our economy there are always more women than men in the marriage market. As a consequence, the probability that a woman meets a younger man, $\theta^w_t = \frac{N_{0,t}}{N_{0,t}+N^m_{1,t}}$, is equivalent to the probability that a man meets a woman. Therefore, the previous result also implies that a positive cohort shock increases the probability that a woman meets a younger men.

We can now determine the effect of a shock to cohort size on the reservation value of younger men $\underline{\eta}_{0,\tau}$. Notice that in the determination of $\underline{\eta}_{0,\tau}$ a younger man compares the value of getting married at τ with the value of waiting until next period. The value of waiting depends on the probability he will meet a woman in period $\tau + 1$. This probability depends on the number of older men at $\tau + 1$, which can be written as follows:

$$\theta^m_{0,\tau} = \theta^m_{1,\tau} = \frac{N_0 + \Delta}{N_0 + \Delta + N_{1,\tau+1}}.$$

Using equation (7), we can substitute for $N_{1,\tau+1}$ to obtain the following expression:

$$\theta^m_{0,\tau} = \theta^m_{1,\tau} = \frac{N_0 + \Delta}{N_0 + \Delta + (N_0 + \Delta)\left(1 - \theta^m_{0,\tau}\left(1 - F\left(\underline{\eta}_{0,\tau}\right)\right)\right)} = \frac{1}{1 + \left(1 - \theta^m_{0,\tau}\left(1 - F\left(\underline{\eta}_{0,\tau}\right)\right)\right)}.$$

We can now substitute for $\theta^m_{1,\tau}$ in the equation that determines $\underline{\eta}_{0,\tau}$ to obtain

$$\underline{\eta}_{0,\tau} = 2\delta + \beta \frac{1 - \beta^T}{1 - \beta^{T+1}} \gamma \left\{ E\left[\eta \,|\eta \geq 2\delta\right] - 2\delta \right\} \left(1 - F\left(2\delta\right)\right) \frac{1}{1 + \left(1 - \theta^m_{0,\tau}\left(1 - F\left(\underline{\eta}_{0,\tau}\right)\right)\right)}. \quad (8)$$

The same equation for the reservation value in steady state can be derived as follows:

$$\underline{\eta}_{0,ss} = 2\delta + \beta \frac{1-\beta^T}{1-\beta^{T+1}} \gamma \left\{ E\left[\eta \left| \eta \geq 2\delta \right.\right] - 2\delta \right\} \left(1 - F\left(2\delta\right)\right) \frac{1}{1 + \left(1 - \theta^m_{0,ss}\left(1 - F\left(\underline{\eta}_{0,ss}\right)\right)\right)}. \quad (9)$$

Earlier in this section we have shown that, with a positive shock to cohort size, $\theta^m_{0,\tau} > \theta^m_{0,ss}$. As a consequence, a simple comparison of the last two equations implies that an increase in cohort size has the effect of increasing the reservation value of young men. This result is summarized in the following proposition.

Proposition 2 *A positive shock to cohort size in period τ increases the reservation value $\underline{\eta}_{0,\tau}$. A negative shock has the opposite effect.*

Proof. In the appendix. ■

Using Proposition 2 we can now show that a positive cohort shock reduces the fraction of women in a given cohort who marry and that a negative shock has the opposite effect. The following Proposition establishes this result.

Proposition 3 *A positive shock to cohort size in period τ reduces the fraction of cohort τ's women who get married by increasing the reservation value of marriage of young couples. A negative shock in period τ has the opposite effect.*

Proof. In the appendix. ■

To provide the insight behind this result, consider an increase in cohort size. After this event, older men become a scarce resource. This change has two effects. First, the fraction of women who marry mechanically declines because more women will not meet a potential spouse. Second, younger men become more selective because they will have a larger group of women to choose from when they will be old. As a consequence, the fraction of women who marry decreases. The total impact of an increase in cohort size is therefore a reduction in the fraction of women who marry. This result indicates that the search model developed in this paper can explain the negative relationship observed in the data between cohort size and marriage rates.

We will now derive an implication of the model that will be used in the next section as a test. In the following Proposition we shows that an increase in cohort size has the effect of decreasing the average difference in age at the time of marriage.

Proposition 4 *In the model, an increase in cohort size reduces the average age difference between spouses. A reduction in cohort size has the opposite effect.*

Proof. In the appendix. ■

The intuition behind Proposition 4 is based on the following two effects. First, women are less likely to meet older men because there are relatively fewer of them. Since the reservation value of older men does not depend on cohort size, a consequence of this is that the average age difference between spouses will decline. Second, women are more likely to meet younger men because of the increase in cohort size. Without the increase in reservation value, this effect would decrease the average age difference at marriage. However, younger men become more selective in their decision to marry. As a result, fewer of the younger men that meet a woman end up getting married. The combined effect of younger men on the average age difference is therefore unclear. Proposition 4 shows that the effect generated by older men and the effect produced by a greater probability of meeting a younger men dominate the effect of a higher reservation value. The increase in cohort size therefore reduces the average age difference at marriage. This result will be used in the next section to test whether the model is consistent with the data.

6 Test and Estimation of the Search Model

This section will be divided into two parts. We will first test the model developed in the previous section using the result contained in Proposition 4. We will then estimate the model with the objective of evaluating whether it can quantitatively match the variation in marriage rates observed in the data. In both parts we will only consider the population of whites because, as argued in the previous section, the assumption $N_0^m = N_0^w = N_0$ is not satisfied in the data for blacks.

6.1 A Test of the Search Model

Proposition 4 establishes an implication of our search model. If the model is correct, an increase in cohort size should reduce the average age difference between spouses, whereas a reduction in cohort size should have the opposite effect. In this subsection, we will use this result to evaluate whether our model is consistent with the patterns observed in the data.

We perform the test in three stages. We first provide some evidence on the relationship between cohort size and average age difference by plotting the time series of these two variables. We then use the time-series variation to regress the logarithm of mean age difference between spouses on the logarithm of cohort size. Finally, we run the same regression using cross-state variation.

To implement the test, we have to construct the variable age difference between spouses. When we use the longitudinal variation, this variable is created using the the CPS years 1962-2010 for cohorts born on or after 1932 and the 1950 Census for cohorts born before 1932. Specifically, for each cohort we consider all women between the ages of 30 and 35 who are married. We then compute the difference between their age and the age of their spouse. Finally, we calculate the average for each cohort. When we employ the cross-state variation, the average age difference is computed using the 1970, 1980, 1990, and 2000 Censuses, since the CPS does not have enough observations at the state level. In this case, we consider all women of age 30 who are married, compute the age difference with their spouse, and calculate the average at the state level. We then determine 10-year differences in the variable of interest at the state level for each one of the cohorts we consider, i.e. for cohorts born in 1940, 1950, 1960, and 1970. Finally, we regress these 10-year differences on the corresponding 10-year differences in cohort size. It should be remarked that in both the longitudinal and cross-state variation the cohort size is measured at the time the individuals in a given cohort are born, whereas the age difference variables is computed three decades later when most of the women in that cohort had made their marriage decisions. It is also important to remark that we cannot perform the test using the mobilization rates. The reason for this is that the age difference cannot be constructed consistently for all cohorts born between 1937 to 1950 using the Censuses, the only data in which we have enough observations at the state level. To compute that variable for the required cohorts, we would have to use the recall variable age at first marriage, which is observed, jointly with a recall variable age of the first spouse, which is not observed. However, this is not an important limitation since testing the implication of the search model does not require exogenous variation.

In Figure 11 we report graphical evidence on the relationship between age differences at marriage and cohort size using longitudinal variation for whites. With the exception of the first twelve cohorts, Figure 11 indicates that there is a tight relationship between age difference at marriage and cohort size. As the model predicts, when the size of a given cohort increases, the age difference between women in that cohort and their spouses becomes less negative and therefore declines. When the cohort size drops, the age difference between spouses becomes more negative and therefore increases.

Table 6 reports the coefficients and R-squared obtained by regressing the logarithm of the average age difference between spouses on the logarithm of cohort size for whites using time-series variation. This regression enables us to determine whether the link between these two variables is statistically significant and how much of the variation in age difference at marriage is explained by cohort size. The estimated coefficient on cohort size is around -0.6 and strongly statistically significant. It indicates that a 10% increase in cohort size generates a decline in age difference at marriage of approximately 6%. The size of the effect is therefore large. Finally, the R-squared suggests that cohort size can explain a significant fraction of the variation across cohorts in age difference between spouses. Our results indicate that about 81% of the variation in this variable can be explained by changes in cohort size. Finally, in the same Table we report the coefficient estimates obtained using cross-state variation. We run three different regressions, one for each of the 10-year differences between cohorts. The estimated coefficients are consistent with the search model. With the exception of the first regression for which our estimated coefficient is statistically equal to zero, our estimates are negative and statistically significant. We therefore cannot reject the model developed in the previous section.

6.2 Estimation of the Search Model

In this subsection, we estimate the dynamic search model developed in this paper with the objective of evaluating whether it can quantitatively explain the changes in marriage rates observed in the data. This exercise is important because it represents an additional and more thorough test of the mechanism behind the relationship between marriage rates and cohort size.

To structurally estimate the model, we have to make additional assumptions. The first assumption is about the distribution of the match quality η. We assume that it is distributed according to a beta distribution with shape parameters α_1 and α_2 defined on the interval $(0, 1)$. We have chosen the beta distribution for two reasons. First, it is one of the most flexible distributions. Evidence of this is that many popular distributions like the uniform, the exponential, and the gamma are special cases of the beta distribution and that the normal distribution can be well approximated by it. The second reason is that the beta distribution is parsimonious with only two parameters to estimate.

A second assumption is required to be able to estimate the model. In the version developed in section 5, there is no source of uncertainty. To address this issue, we assume that the value of being single δ varies over time according to the following equation:

$$\delta_t = \delta + \nu_t,$$

where ν_t is drawn from a uniform distribution defined on the interval $[-0.2, 0.2]$.[5]

The third set of assumptions we make are related to the lifespan of an individual. Observe that individuals born in a given cohort marry over many years. Some of them find a spouse the first time they enter the marriage market, whereas others marry after having searched for many years. This implies that, in any given year, individuals from different cohorts compete in the marriage market. To model this feature, we assume that each period in our model corresponds to 10 years of an individual's life, that an individual starts making decisions at age 20, and that she or he lives for 50 years or, equivalently, 5 periods. To implement the assumption that each

[5] Changing the interval does not change the outcome of the estimation.

period corresponds to ten years, in a period we allow each individual to meet sequentially as many as ten potential spouses, one for each year. The individual leaves the marriage market when she or he marries one of the potential spouses. With this additional feature, an increases in cohort size that follows a previous rise will have a larger effect than a single increase, because the newcomers compete in a marriage market that is already crowded by the first rise. A similar argument applies to declines in cohort size.[6]

As additional assumptions, we set the annual discount factor equal to 0.98 and consider a symmetric Nash-bargaining by setting γ equal to 0.5. Finally, we augment the model to allow for a fraction of men that are unwilling to marry independently of the value of match quality. We will denote this fraction with $1-\phi$. This parameter only affects the probability that a younger or older man meets a woman. Specifically, these probabilities now take the following form:

$$\theta_{0,t}^{m} = \theta_{1,t}^{m} = \frac{N_{0,t}^{w}}{\left(N_{0,t}^{m} + N_{1,t}^{m}\right)\phi}.$$

The probability that a woman meets a younger or older man does not change since the parameter ϕ appears at the numerator as well as denominator.

Given these assumptions, the model has four parameters that must be estimated: the value of being single δ, which is assumed to be identical across gender and over time; the two shape parameters of the beta distribution α_1 and α_2; the fraction of individuals that are unwilling to marry $1-\phi$. These parameters will be estimated using Simulated Method of Moments (McFadden (1989), Pakes and Pollard (1989), Lee and Ingram (1991), and Duffie and Singleton (1993)). Specifically, the estimation is performed in three steps. For a given set of parameters that characterize the model, we first simulate the individual decisions. We then compute a function of the differences between some of the statistical moments that characterize the data and the corresponding moments obtained from the simulated data. Finally, the estimated parameters are obtained by minimizing this function.

[6]We have decided not to include this feature in the theory part because it makes the model more complicated without changing the insight it provides.

In the estimation, we use as our set of moments the fraction of women never married in a cohort starting from the cohort born in 1930 and ending with the cohort born in 1980. We use cohorts born from 1909 to 1929 to initialize the model. We therefore have 51 moments that will be matched using 4 parameters. Before presenting the results it is important to remark that these moments enable us to identify the value of being single δ and the parameter that determines the fraction of men who are unwilling to marry ϕ. To see this observe that the value of being single δ is linked to the fraction of individuals never married in a given cohort. Everything else equal, a higher value of δ increases the share of individuals who choose to stay single in each cohort. The parameter ϕ enters equation (5) which defines the reservation value of young men. That equation can be viewed as a linear relationship between the reservation value of young men and the probability that when old they will meet a potential spouse. The slope of that equation is affected by ϕ: a larger value for ϕ generates a smaller slope. As a consequence, changes in the meeting probabilities will have smaller effects on the reservation value of younger men and therefore on marriage rates if ϕ is larger. In the model, changes in meeting probabilities are mainly generated by variations in cohort size. The parameter ϕ can, therefore, be identified by measuring how the fraction of never married individuals varies in response to changes in cohort size. However, there is no reason to believe that the remaining two parameters can be identified using the selected set of moments. As consequence, the exercise performed in this subsection should be seen as a test of whether there exist parameter values for δ, α_1, α_2, and ϕ such that the dynamic search model can quantitatively explain the variation in marriage rates observed in the data. We believe that this exercise is the most interesting to perform, since in this paper we are not interested in performing policy evaluation or predictions.[7]

The estimated parameters are reported in Table 7. Our estimates of α_1 and α_2 are 0.020 and 0.072. These values imply that the mean of the distribution is equal to 0.217, the mode is equal

[7] We have proven that the parameter α_1 can be identified using as a moment the probability that an older man marries plus the corresponding probability for a woman; α_2 can be identified using as a moment the probability that a younger man marries divided by the probability that a woman marries times the previous moment; the parameter ϕ can be identified using as a moment the probability that a woman does not marry divided by the sum of the probability that a young man does not marry and of the probability that an older man does not marry; the parameter γ can be identified using the fact that it changes the slope of the equation characterizing the reservation value. We will use these moments for the estimation of the parameters in future research.

to 0.512, and the standard deviation is equal to 0.036. The value of being single is estimated to be 0.107. This value indicates that only couples with relatively high match quality will choose to marry. To see this, remember that a woman and an older man marry only if $\eta > 2\delta$. Given our estimate of δ, this means that couples will decide to marry only if the drawn match quality value is higher than 0.214. Since, this number is approximately equal to the mean, women matched to older men marry only if they draw a relatively high value for match quality. Finally, observe that younger men have a higher reservation value. Consequently, women paired with younger men will also marry only if their match quality is relatively high. Finally, our model rationalized the data by estimating that 13.3% of men are unwilling to marry independently of match quality.

We will now evaluate whether the model can quantitatively match the relationship between changes in cohort size and changes in marriage rates. We do this by plotting in Figure 12 the marriage rates observed in the data jointly with the marriage rates simulated using the estimated model, both as a function of cohort. Figure 12 shows that the model can quantitatively replicate the variation in the share of never married across cohorts observed in the data. This outcome is particularly remarkable since our model is very parsimonious. We only have four parameters to match 51 moments.

7 Conclusions

In this paper we provide an explanation for the variation in U.S. marriage rates over the past century. Using time-series variation, cross-state variation, and variation in differences across states in mobilization rates during World War II we provide evidence in support of the following two results. First, cohort size can explain on its own almost all the variation in U.S. marriage rates. Second, an increase in cohort size reduces marriage rates and a reduction has the opposite effect. We then develop a dynamic search model of the marriage market that has the potential of explaining the patterns observed in the data. Using the model, we first show that qualitatively it can generate the relationship between cohort size and marriage rates. We then derive the following testable implication: in the model an increase in cohort size reduces the age difference between spouses and vice versa. Finally, we test the model in two different ways. We first test whether the derived implication can be rejected. In the data, a rise in cohort size reduces the age difference at marriage and a decline in cohort size has the opposite effect. We therefore cannot reject our search model. We then estimate the model and evaluate whether it can quantitatively match the link between cohort size and marriage rates we document. The estimated model can easily match most of the variation in marriage rates observed in the data.

References

Abramitzky, Ran, Adeline Delavande, and Luis Vasconcelos. 2011. "Marrying Up: The Role of Sex Ratio in Assortative Matching." *American Economic Journal: Applied Economics* 3 (3): pp. 124–157.

Acemoglu, Daron, David H. Autor, and David Lyle. 2004. "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury." *Journal of Political Economy* 112 (3): pp. 497–551.

Akerlof, George A., Janet L. Yellen, and Michael L. Katz. 1996. "An Analysis of Out-of-Wedlock Childbearing in the United States." *The Quarterly Journal of Economics* 111 (2): pp. 277–317.

Angrist, Josh. 2002. "How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation." *The Quarterly Journal of Economics* 117 (3): pp. 997–1038.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy* 81 (4): pp. 813–846.

———. 1974. "A Theory of Marriage: Part II." *Journal of Political Economy* 82 (2): pp. S11–S26.

Bergstrom, Theodore C., and Mark Bagnoli. 1993. "Courtship as a Waiting Game." *Journal of Political Economy* 101 (1): pp. 185–202.

Bitler, Marianne, and Lucie Schmidt. 2011. "Marriage Markets and Family Formation: The Role of the Vietnam Draft." *University of California at Irvine Working Paper.*

Charles, Kerwin Kofi, and Ming Ching Luoh. 2010. "Male Incarceration, the Marriage Market, and Female Outcomes." *The Review of Economics and Statistics* 92 (3): pp. 614–627.

Chiappori, Pierre-Andr, Murat Iyigun, and Yoram Weiss. 2009. "Investment in Schooling and the Marriage Market." *The American Economic Review* 99 (5): pp. 1689–1713.

Choo, Eugene, and Aloysius Siow. 2006. "Who Marries Whom and Why." *Journal of Political Economy* 114 (1): pp. 175–201.

Duffie, Darrell, and Kenneth J. Singleton. 1993. "Simulated Moments Estimation of Markov Models of Asset Prices." *Econometrica* 61 (4): pp. 929–952.

Easterlin, Richard A. 1987. *Birth and Fortune: The Impact of Numbers on Personal Welfare.* Chicago: Chicago University Press.

Ellwood, David T., and Jonathan Crane. 1990. "Family Change Among Black Americans: What Do We Know?" *The Journal of Economic Perspectives* 4 (4): pp. 65–84.

Gale, D., and L. S. Shapley. 1962. "College Admissions and the Stability of Marriage." *The American Mathematical Monthly* 69 (1): pp. 9–15.

Greenwood, Jeremy, and Nezih Guner. 2009. "Marriage and Divorce since World War II: Analyzing the Role of Technological Progress on the Formation of Households." *NBER Macroeconomics Annual* 23 (231-276): pp.

Grossbard-Shechtman, Amyra. 1984. "A Theory of Allocation of Time in Markets for Labour and Marriage." *The Economic Journal* 94 (376): pp. 863–882.

Guttentag, Marcia, and Paul F. Secord. 1983. *Too many women?: the sex ratio question.* Beverly Hills, CA: Sage Publications.

Hill, Matthew. 2011. "Love in the Time of Depression: The Effect of Economic Downturns on the Probability of Marriage." *UCLA Working Paper.*

Hitsch, Gunter J. 2010. "Matching and Sorting in Online Dating." *American Economic Review* 100 (1): 130–63 (March).

Iyigun, Murat, and Randall P. Walsh. 2007. "Building the Family Nest: Premarital Investments, Marriage Markets, and Spousal Allocations." *The Review of Economic Studies* 74 (2): pp. 507–535.

Lee, Bong-Soo, and Beth Fisher Ingram. 1991. "Simulation Estimation of Time-series mMdels." *Journal of Econometrics* 47 (2-3): pp. 197–205.

Lerman, Robert I. 1989. "Employment Opportunities of Young Men and Family Formation." *The American Economic Review, Papers and Proceedings* 79 (2): pp. 62–66.

McFadden, Daniel. 1989. "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration." *Econometrica* 57 (5): pp. 995–1026.

Mortensen, Dale T. 1988. "Matching: Finding a Partner for Life or Otherwise." *American Journal of Sociology* 94:pp. S215–S240.

Pakes, Ariel, and David Pollard. 1989. "Simulation and the Asymptotics of Optimization Estimators." *Econometrica* 57 (5): pp. 1027–1057.

Pampel, Fred C., and H. Elizabeth Peters. 1995. "The Easterlin Effect." *Annual Review of Sociology* 21:pp. 163–194.

Peters, Michael, and Aloysius Siow. 2002. "Competing Premarital Investments." *Journal of Political Economy* 110 (3): pp. 592–608.

Plotnick, Robert D. 1990. "Welfare and Out-of-Wedlock Childbearing: Evidence from the 1980s." *Journal of Marriage and Family* 52 (3): pp. 735–746.

Schoen, Robert. 1983. "Measuring the Tightness of a Marriage Squeeze." *Demography* 20 (1): pp. 61–78.

Wilson, William J. 1987. *The truly disadvantaged: the inner city, the underclass, and public policy.* Chicago: Chicago University Press.

Wolfers, Justing. 2010. "What Is Going on With Marriage?" *New York Times: Op-Ed Contribution and* $http://www.freakonomics.com/2010/10/13/what-is-going-on-with-marriage/$*, October 12.*

A Proofs

A.1 Proof of Proposition 2

Consider a positive change to cohort size. According to equation (9), in steady state the reservation value of a young man is the solution to the following equation:

$$\underline{\eta}_{0,ss} = 2\delta + \beta \frac{1-\beta^T}{1-\beta^{T+1}} \gamma \left\{ E\left[\eta \left| \eta \geq 2\delta \right.\right] - 2\delta \right\} \left(1 - F\left(2\delta\right)\right) \frac{1}{1 + \left(1 - \theta^m_{0,ss}\left(1 - F\left(\underline{\eta}_{0,ss}\right)\right)\right)}.$$

By substituting $\theta^m_{0,ss}$ with $\theta^m_{0,\tau}$ and by using the result that $\theta^m_{0,\tau} > \theta^m_{0,ss}$, we obtain the following inequality:

$$\underline{\eta}_{0,ss} < 2\delta + \beta \frac{1-\beta^T}{1-\beta^{T+1}} \gamma \left\{ E\left[\eta \left| \eta \geq 2\delta \right.\right] - 2\delta \right\} \left(1 - F\left(2\delta\right)\right) \frac{1}{1 + \left(1 - \theta^m_{0,\tau}\left(1 - F\left(\underline{\eta}_{0,ss}\right)\right)\right)}.$$

Since the left hand side of the inequality is increasing in $\underline{\eta}_0$ and the right hand side is decreasing in $\underline{\eta}_0$, equation (8) implies that $\underline{\eta}_{0,\tau} > \underline{\eta}_{0,ss}$.

A.2 Proof of Proposition 3

The total number of women that marry in a particular cohort is given by the total number of women in the cohort time the probability that a woman in that cohort marries. As a consequence, the fraction of women in a cohort that marries is simply the probability of marriage for those women. The probability that a woman marries can be written as the probability that she meets a younger man times the probability she marries him plus the probability she meets an older man times the probability she marries him, i.e.

$$P\left(\text{woman marries at } \tau\right) = \theta^w_{0,\tau}\left(1 - F\left(\underline{\eta}_{0,\tau}\right)\right) + \left(1 - \theta^w_{0,\tau}\right)\left(1 - F\left(2\delta\right)\right)$$

Define $1+\lambda_\tau = \frac{F\left(\underline{\eta}_{0,\tau}\right)}{F\left(\underline{\eta}_{0,ss}\right)}$ and $1+\phi_\tau = \frac{\theta^w_{0,\tau}}{\theta^w_{0,ss}}$, where $\lambda_\tau > 0$ and $\phi_\tau > 0$ because $\frac{\partial \underline{\eta}_{0,\tau}}{\partial N_0} > 0$ and $\frac{\partial \theta^w_{0,\tau}}{\partial N_0} > 0$. We then have

$$
\begin{aligned}
& P\left(\text{woman marries at } \tau\right) = \\
= \quad & \theta^w_{0,\tau}\left(1 - F\left(\underline{\eta}_{0,\tau}\right)\right) + \left(1-\theta^w_{0,\tau}\right)\left(1 - F\left(2\delta\right)\right) \\
= \quad & \theta^w_{0,ss}\left(1+\phi_\tau\right)\left(1 - F\left(\underline{\eta}_{0,ss}\right)\left(1+\lambda_\tau\right)\right) + \left(1-\theta^w_{0,ss}\left(1+\phi_\tau\right)\right)\left(1 - F\left(2\delta\right)\right) \\
= \quad & \theta^w_{0,ss}\left(1 - F\left(\underline{\eta}_{0,ss}\right)\right) + \left(1-\theta^w_{0,ss}\right)\left(1 - F\left(2\delta\right)\right) - \theta^w_{0,ss}\lambda_\tau F\left(\underline{\eta}_{0,ss}\right) + \theta^w_{0,ss}\phi_\tau\left(1 - F\left(\underline{\eta}_{0,ss}\right)\left(1+\lambda_\tau\right)\right) \\
& - \theta^w_{0,ss}\phi_\tau\left(1 - F\left(2\delta\right)\right) \\
= \quad & P\left(\text{woman marries at } ss\right) - \theta^w_{0,ss}\lambda_\tau F\left(\underline{\eta}_{0,ss}\right) + \theta^w_{0,ss}\phi_\tau\left(1 - F\left(\underline{\eta}_{0,\tau}\right)\right) - \theta^w_{0,ss}\phi_\tau\left(1 - F\left(2\delta\right)\right) \\
< \quad & P\left(\text{woman marries at } ss\right) - \theta^w_{0,ss}\lambda_\tau F\left(\underline{\eta}_{0,ss}\right) \\
< \quad & P\left(\text{woman marries at } ss\right).
\end{aligned}
$$

A.3 Proof of Proposition 4

The average difference in age between spouses at marriage can be computed as the difference in age conditional on the woman marrying a younger man times the corresponding probability plus the difference in age conditional on marrying an older man times the corresponding probability, i.e.

$$E\left[\Delta age\right] = E\left[\Delta age | younger\ man\right] P\left(younger\ man\right) + E\left[\Delta age | older\ man\right] P\left(older\ man\right).$$

Without loss of generality suppose the difference in age if a woman marries a younger man is equal to y_1 whereas the corresponding difference in a marriage with an old man is equal to y_2 with $y_1 < y_2 = y_1 + z$. Let $P = \theta^w_{0,\tau}\left(1 - F\left(\underline{\eta}_{0,\tau}\right)\right) + \left(1-\theta^w_{0,\tau}\right)\left(1 - F\left(2\delta\right)\right)$ be the probability that a woman marries and let $x = \underline{\eta}_{0,\tau}$. Then, $E[\Delta age]$ can be written as follows:

$$
\begin{aligned}
E[\Delta age] &= \frac{y_1 \theta^w_{0,\tau}(1-F(x)) + y_2\left(1-\theta^w_{0,\tau}\right)(1-F(2\delta))}{P} \\
&= y_1 + \frac{z\dfrac{N_1}{N_0+N_1}(1-F(2\delta))}{P} \\
&= y_1 + \frac{z}{\dfrac{N_0(1-F(x))}{N_1(1-F(2\delta))}+1}.
\end{aligned}
$$

Notice that y_1, z, and $1-F(2\delta)$ are constant with respect to changes in cohort size N_0. In addition, N_1 is also constant with respect to N_0 since it measures the number of older men in the marriage market the period of the shock which is not affected by contemporaneous cohort size.

To prove the proposition, we have therefore to prove that an increase in N_0 increases $N_0(1-F(x))$. We will prove this result by contradiction. Consider a small change in cohort size Δ and suppose it reduces $N_0(1-F(x))$. We will show that, if this is true, the number of older men in the period that follows the shock to cohort size is so large that the reservation value of younger men in the period of the shock is smaller than the reservation value in steady state. This result contradicts the previous finding that an increase in N_0 increases x. Hence, an increase in N_0 must reduce $N_0(1-F(x))$.

Denote with θ' and x' the probability that a men meets a woman and the reservation value of a younger men both in the period in which the increase in N_0 occurs. Furthermore, denote with N_1' the number of old men the period that follows the increase in N_0. The change in N_1' can be computed as the number of younger men who in this period do not meet a woman plus the number of younger men who meet a woman and decide not to marry. Hence,

$$\begin{aligned}
N_1' - N_1 &= (N_0 + \Delta)\left(1 - \theta'\right) + (N_0 + \Delta)\,\theta' F\left(x'\right) - N_0\,(1 - \theta) - N_0 \theta F\,(x) \\
&= -\,(N_0 + \Delta)\,\theta'\left(1 - F\left(x'\right)\right) + N_0 \theta\,(1 - F\,(x)) + N_0 + \Delta - N_0 \\
&= -\,(N_0 + \Delta)\,\theta'\left(1 - F\left(x'\right)\right) + N_0 \theta'\,(1 - F\,(x)) + N_0\left(\theta - \theta'\right)(1 - F\,(x)) + \Delta \\
&= -\theta'\left((N_0 + \Delta)\left(1 - F\left(x'\right)\right) - N_0\,(1 - F\,(x))\right) + N_0\left(\theta - \theta'\right)(1 - F\,(x)) + \Delta \\
&\geq N_0\left(\theta - \theta'\right)(1 - F\,(x)) + \Delta, \qquad (10)
\end{aligned}$$

where the inequality follows from the assumption that $(N_0 + \Delta)\,(1 - F\,(x')) - N_0\,(1 - F\,(x)) < 0$.

We will now show that $N_1' - N_1 \geq N_0\,(\theta - \theta')\,(1 - F\,(x)) + \Delta$ implies that θ'', the probability that a man meets a woman the period after the shock takes place, is lower than θ. As a consequence, the reservation value in the period of the shock is lower than before the shock. Observe that

$$\theta = \frac{N_0}{N_0 + N_1} = \frac{1}{1 + \dfrac{N_1}{N_0}}$$

and

$$\theta'' = \frac{N_0 + \Delta}{N_0 + \Delta + N_1'} = \frac{1}{1 + \dfrac{N_1'}{N_0 + \Delta}}.$$

Hence, $\theta > \theta''$ if and only if $\dfrac{N_1'}{N_0 + \Delta} > \dfrac{N_1}{N_0} = \sqrt{F\,(x)}$. Equation (10) implies that

$$\frac{N_1'}{N_0 + \Delta} \geq \frac{N_1 + N_0\,(\theta - \theta')\,(1 - F\,(x)) + \Delta}{N_0 + \Delta}.$$

Simple algebra implies that $\dfrac{N_1 + N_0\,(\theta - \theta')\,(1 - F\,(x)) + \Delta}{N_0 + \Delta} > \dfrac{N_1}{N_0}$ if

$$\frac{N_0\,(\theta - \theta')\,(1 - F\,(x)) + \Delta}{\Delta} > \frac{N_1}{N_0} = \sqrt{F\,(x)}.$$

Replacing for θ and θ', the left hand side of the inequality can be written as follows:

$$
\begin{aligned}
& \frac{N_0\left(1-F\left(x\right)\right)}{\Delta}\left(\frac{1}{1+\sqrt{F\left(x\right)}}-\frac{N_0+\Delta}{N_0+\Delta+N_0\sqrt{F\left(x\right)}}\right)+1 \\
= \quad & \frac{N_0\left(1-F\left(x\right)\right)}{\Delta}\left(\frac{N_0+\Delta+N_0\sqrt{F\left(x\right)}-\left(N_0+\Delta\right)\left(1+\sqrt{F\left(x\right)}\right)}{\left(1+\sqrt{F\left(x\right)}\right)\left(N_0+\Delta+N_0\sqrt{F\left(x\right)}\right)}\right)+1 \\
= \quad & \frac{N_0\left(1-F\left(x\right)\right)}{\Delta}\frac{-\Delta\sqrt{F\left(x\right)}}{\left(1+\sqrt{F\left(x\right)}\right)\left(N_0+\Delta+N_0\sqrt{F\left(x\right)}\right)}+1 \\
= \quad & \frac{-N_0\left(1-F\left(x\right)\right)\sqrt{F\left(x\right)}+\left(1+\sqrt{F\left(x\right)}\right)\left(N_0+\Delta+N_0\sqrt{F\left(x\right)}\right)}{\left(1+\sqrt{F\left(x\right)}\right)\left(N_0+\Delta+N_0\sqrt{F\left(x\right)}\right)} \\
= \quad & \frac{N_0+\Delta+N_0\sqrt{F\left(x\right)}+\Delta\sqrt{F\left(x\right)}+N_0F\left(x\right)+N_0F\left(x\right)\sqrt{F\left(x\right)}}{N_0+\Delta+2N_0\sqrt{F\left(x\right)}+\Delta\sqrt{F\left(x\right)}+N_0F\left(x\right)}
\end{aligned}
$$

Hence, we have the desired inequality if

$$
\frac{N_0+\Delta+N_0\sqrt{F\left(x\right)}+\Delta\sqrt{F\left(x\right)}+N_0F\left(x\right)+N_0F\left(x\right)\sqrt{F\left(x\right)}}{N_0+\Delta+2N_0\sqrt{F\left(x\right)}+\Delta\sqrt{F\left(x\right)}+N_0F\left(x\right)}>\sqrt{F\left(x\right)},
$$

or equivalently,

$$
\begin{aligned}
& N_0+\Delta+N_0\sqrt{F\left(x\right)}+\Delta\sqrt{F\left(x\right)}+N_0F\left(x\right)+N_0F\left(x\right)\sqrt{F\left(x\right)} \\
> \quad & N_0\sqrt{F\left(x\right)}+\Delta\sqrt{F\left(x\right)}+2N_0F\left(x\right)+\Delta F\left(x\right)+N_0F\left(x\right)\sqrt{F\left(x\right)}.
\end{aligned}
$$

Some of the terms cancel out producing the following inequality:

$$
N_0+\Delta>N_0F\left(x\right)+\Delta F\left(x\right),
$$

which is equivalent to

$$
1>F\left(x\right),
$$

which is always satisfied. As a consequence, $\theta>\theta''$ which implies that the reservation value of young men in steady state is greater than their reservation value in the period of the shock, which contradict our result that the reservation value increases with an increase in cohort size. As a consequence, $N_0 (1 - F(x))$ must increase with cohort size. Hence, the expected value of the age difference at marriage declines with a positive shock to cohort size.

A.4 Expected Value Functions

For completeness, in this appendix we derive the expected values for young men and women. The expected value of a young man takes the following form:

$$
\begin{aligned}
v_{0,t}^m &= \theta_{0,t}^m \left(1 - F\left(\underline{\eta}_{1,t}\right)\right) \left\{\delta + \beta v_{1,t}^m + \gamma \left\{\frac{1-\beta^{T+1}}{1-\beta} E\left[\eta \left| \eta \geq \underline{\eta}_{0,t}\right.\right] - \left(\delta + \beta v_{1,t}^m\right) - \frac{1-\beta^{T+1}}{1-\beta}\delta\right\}\right\} \\
&+ \theta_{0,t}^m F\left(\underline{\eta}_{0,t}\right)\left(\delta + \beta v_{1,t}^m\right) + \left(1 - \theta_{0,t}^m\right)\left(\delta + \beta v_{1,t}^m\right).
\end{aligned}
$$

The first term represents the value of meeting a woman with a match quality η higher than the reservation value times the probability of this event. The second term describes the value of meeting a women characterized by an η lower than the reservation value multiplied by the corresponding probability. The third term measures the value of not meeting a woman when young times the probability.

To derive the woman's expected value function we have to take into account that she can meet both younger and older men. As a consequence, it takes the following more complex form:

$$
\begin{aligned}
v_{0,t}^w &= \theta_{0,t}^m \left(1 - F\left(\underline{\eta}_{0,t}\right)\right) \left\{\frac{1-\beta^{T+1}}{1-\beta}\delta + (1-\gamma) \left\{\frac{1-\beta^{T+1}}{1-\beta} E\left[\eta \left| \eta \geq \underline{\eta}_{0,t}\right.\right] - \left(\delta + \beta v_{1,t}^m\right) - \frac{1-\beta^{T+1}}{1-\beta}\delta\right\}\right\} \\
&+ \theta_{0,t}^m F\left(\underline{\eta}_{0,t}\right) \frac{1-\beta^{T+1}}{1-\beta}\delta \\
&+ \theta_{1,t}^m \left(1 - F(2\delta)\right) \left\{\frac{1-\beta^{T+1}}{1-\beta}\delta + (1-\gamma) \left\{\frac{1-\beta^{T}}{1-\beta} E\left[\eta \left| \eta \geq 2\delta\right.\right] + \beta^T \delta - v_{1,t}^m - \frac{1-\beta^{T+1}}{1-\beta}\delta\right\}\right\} \\
&+ \theta_{1,t}^m F(2\delta) \frac{1-\beta^{T+1}}{1-\beta}\delta
\end{aligned}
$$

The first term measures the value of meeting a younger man with an η higher than the reservation value times the corresponding probability. The second term is the value of meeting a younger men whom it is optimal not to marry times the probability of this event. The third and fourth terms describe the same values for older men.

B Tables and Figures

Table 1: Time Series Regression of Log Share Ever Married on Log Cohort Size

	White Men	Black Men	White Women	Black Women
Ever married	-0.467*	-0.966*	-0.365*	-1.230*
by age 25	(0.107)	(0.102)	(0.075)	(0.113)
R^2	0.21	0.56	0.25	0.63
Ever married	-0.294*	-0.592*	-0.193*	-0.870*
by age 30	(0.037)	(0.048)	(0.026)	(0.064)
R^2	0.50	0.70	0.46	0.74
Ever married	-0.182*	-0.440*	-0.111*	-0.560*
by age 35	(0.015)	(0.031)	(0.011)	(0.043)
R^2	0.71	0.77	0.65	0.74
Ever married	-0.107*	-0.322*	-0.066*	-0.453*
by age 40	(0.008)	(0.019)	(0.007)	(0.026)
R^2	0.77	0.84	0.58	0.85

* Significant at 1%. Standard errors in parentheses.
Notes: Each coefficient is the outcome of a separate regression. "Share ever married" is the share ever married or currently cohabiting by age a. Data comes from IPUMS CPS 1962-2011, IPUMS Census 1960-1980. Regressions are for cohorts born after 1914 until the most recent cohort observed at age a in 2011.

Table 2: Cross-Sectional Regression of Log Share Ever Married by 30 or 40

A. Dependent Variable: 10-Yr. Difference in Log Share Ever Married by 30

	White Men	Black Men	White Women	Black Women
10-Yr. Difference in	-0.080**	-0.090**	-0.064**	-0.092**
Log Cohort Size	(0.028)	(0.024)	(0.017)	(0.019)
N	144	112	144	112
R^2	0.42	0.74	0.23	0.53

B. Dependent Variable: 10-Yr. Difference in Log Share Ever Married by 40

	White Men	Black Men	White Women	Black Women
10-Yr. Difference in	-0.037**	-0.049*	-0.037**	-0.077**
Log Cohort Size	(0.007)	(0.020)	(0.007)	(0.014)
N	96	74	96	74
R^2	0.30	0.36	0.12	0.51

* Significant at 10%. ** Significant at 1%. Standard errors in parentheses.
Notes: Each coefficient is the outcome of a separate regression. We control for time fixed effects. Regressions include Washington, D.C., and all states except Hawaii, Alaska, and those states with less than 500 births in a particular year. "Share ever married" is the share ever married or currently cohabiting. Data comes from IPUMS Census 1970-2000, for decennial cohorts born starting in 1940. Regressions in Panel B exclude the 1970 cohort, as 30 is the maximum age we can observe for this cohort in the 2000 Census.

Table 3: Summary Statistics, 1940: Low, Medium, and High-Mobilization States (White Only)

	Low	Medium	High
Percent Men Inducted into Army	0.28	0.30	0.32**
	(0.01)	(0.01)	(0.01)
Share Never Married at 30	0.24	0.22	0.21
	(0.06)	(0.05)	(0.05)
Share Farmers	0.24	0.20	0.16
	(0.13)	(0.09)	(0.14)
Age	34.84	34.34	34.15*
	(0.79)	(1.19)	(0.92)
Men's Employment	0.85	0.84	0.83
	(0.02)	(0.03)	(0.03)
Women's Employment	0.26	0.27	0.28
	(0.08)	(0.05)	(0.07)
Log Income	6.55	6.59	6.58
	(0.25)	(0.18)	(0.31)
Years of Education	9.91	9.61	9.16**
	(0.62)	(0.70)	(0.53)
Number of Children by Age 35	1.74	2.00	1.90
	(0.46)	(0.38)	(0.46)

* Difference in means between high- and low-mobilization groups significant at 5%. ** Significant at 1%.
Notes: Standard errors in parentheses. Differences between groups are not significant at the 10%-level unless otherwise noted. Averages are for white individuals only. Total number of observations equals 49. Low-mobilization states (bottom third of states) are those that report the share of men inducted as less than or equal to 0.288. High-mobilization states (top third) are those where the share of men inducted into the army is exceeds 0.312. Analysis includes 48 states (all except Hawaii and Alaska), as well as Washington, D.C. Percent inducted is the cumulative number of white men ages 18-44 inducted, divided by the number of white men ages 18-44 registered, as of Sept. 1, 1945. Numbers taken from the Selective Service System's Special Monograph no. 12, Appendix F, Table 164. All other averages constructed using IPUMS Census, 1940. Share never married at 30 is calculated for white men ages 30 and 31. Observations used to construct the average number of children by age 35 are all white women ages 35 and 36. Averages for these two variables exclude Nevada due to small sample size in the state under the age restrictions. None of the results are substantively affected by including or excluding Nevada. Share farmers is the share of white men ages 18 to 50 with a farming occupation. Age and log income are reported for white men ages 18 to 50. Employment and years of education are calculated for white men and women ages 18 to 50.

Table 4: First-Stage Regression: Log Births and Mobilization

	(1)	(2)	(3)	(4)
Mobilization * 1941	0.227	0.042	0.345	0.291
	(0.483)	(0.505)	(0.545)	(0.545)
Mobilization * 1942	0.681	0.617	1.193**	1.175**
	(0.492)	(0.495)	(0.557)	(0.555)
Mobilization * 1943	-0.151	0.351	0.794	1.098
	(0.500)	(0.694)	(0.568)	(0.733)
Mobilization * 1944	-1.259**	1.528	0.004	1.832
	(0.502)	(1.183)	(0.570)	(1.198)
Mobilization * 1945	-1.545***	-1.584***	-0.606	-0.617
	(0.497)	(0.495)	(0.564)	(0.561)
Mobilization * 1946	-0.835*	-0.874*	-0.061	-0.084
	(0.491)	(0.489)	(0.555)	(0.553)
Mobilization * 1947	-0.604	-0.623	0.033	0.025
	(0.487)	(0.485)	(0.550)	(0.548)
Year-by-Year Mobilization * 1941		2.973		1.957
		(2.833)		(2.740)
Year-by-Year Mobilization * 1942		0.466		-0.154
		(1.028)		(1.174)
Year-by-Year Mobilization * 1943		-0.818		-0.601
		(0.724)		(0.847)
Year-by-Year Mobilization * 1944		-2.897***		-1.990*
		(1.111)		(1.150)
Casualty Rate		-2.429		-5.706**
		(2.408)		(2.428)

* Significant at 10%. ** Significant at 5%. *** Significant at 1%. Standard errors in parentheses.
Notes: The variable "Mobilization Rate * Year " is the interaction between the share of white men ages 18 to 44 inducted into the armed forces (see note in Table 2) and a year dummy. "Year-by-Year Mobilization * Year" is constructed similarly but uses instead the share inducted into the armed forces up the specified year. Columns 1 and 2 allow for year-fixed effects, while columns 3 and 4 allow for time-region fixed effects. Regressions additionally include state fixed effects, as well as controls for 1940 baseline years of education, log income, and share farmers, interacted with a fourth-order polynomial in time.

Table 5: Second Stage Regression: Log Marriage Rates and Log Cohort Size

		(1)	(2)	(3)	(4)
(Men)	Log Cohort Size	-0.027*	-0.053*	-0.027*	-0.052*
		(0.003)	(0.009)	(0.003)	(0.009)
(Women)	Log Cohort Size	-0.023*	-0.029*	-0.023*	-0.029*
		(0.001)	(0.006)	(0.001)	(0.006)

* Significant at 1%. Standard errors in parenthesis. Robust standard errors are clustered at the state-level. Each coefficient is the outcome of a separate regression. The dependent variable is the share of men ever married by 30 or the share of women ever married by 30. Columns 1 and 2 allow for year-fixed effects, while columns 3 and 4 allow for time-region fixed effects. Regressions additionally include state fixed effects, as well as controls for 1940 baseline years of education, log income, and share farmers interacted with a fourth-order polynomial in time.

Table 6: Regressions: Log Age Difference and Log Cohort Size

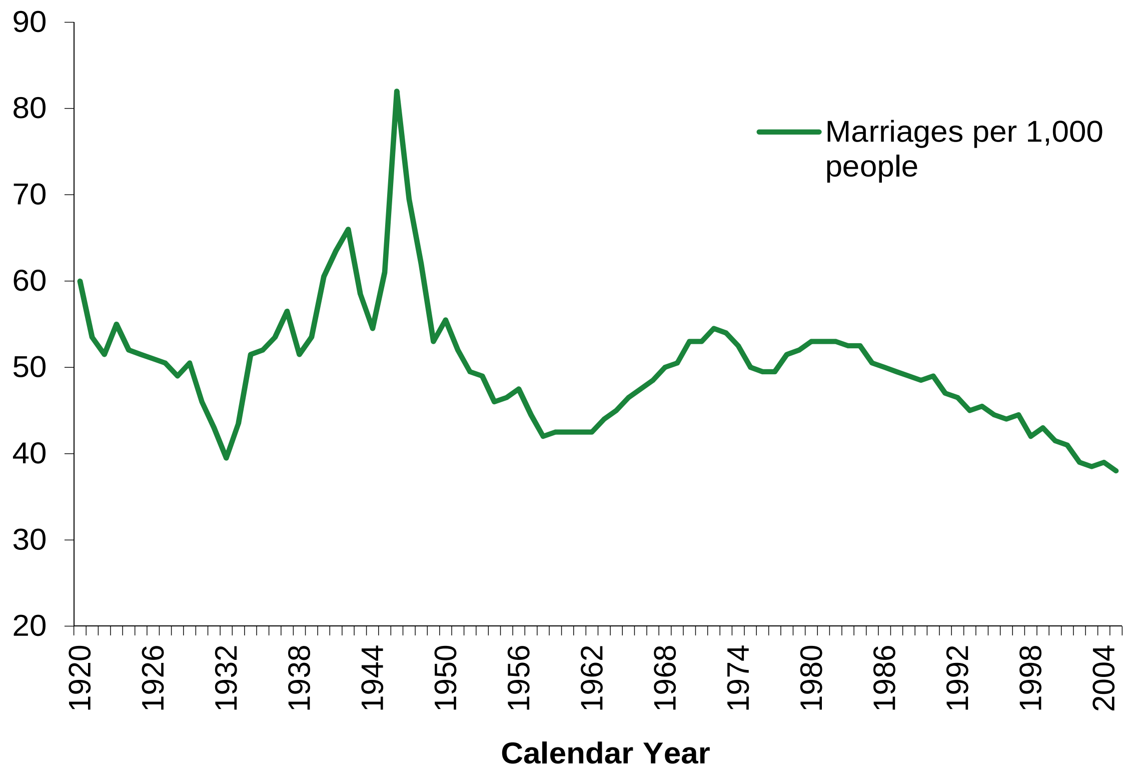
	Time Series, 1930-1975	Cross-State, 1940-1950	Cross-State, 1950-1960	Cross-State, 1960-1970
Log Cohort Size	-0.592***			
	(0.043)			
10-Yr. Difference in		0.110	-0.172**	-0.130*
Log Cohort Size		(0.235)	(0.077)	(0.078)
N	46	49	49	49
R-squared	0.81	0.00	0.05	0.09

* Significant at 10%. ** Significant at 5%. *** Significant at 1%. Standard errors in parentheses.

Table 7: Estimated Parameters

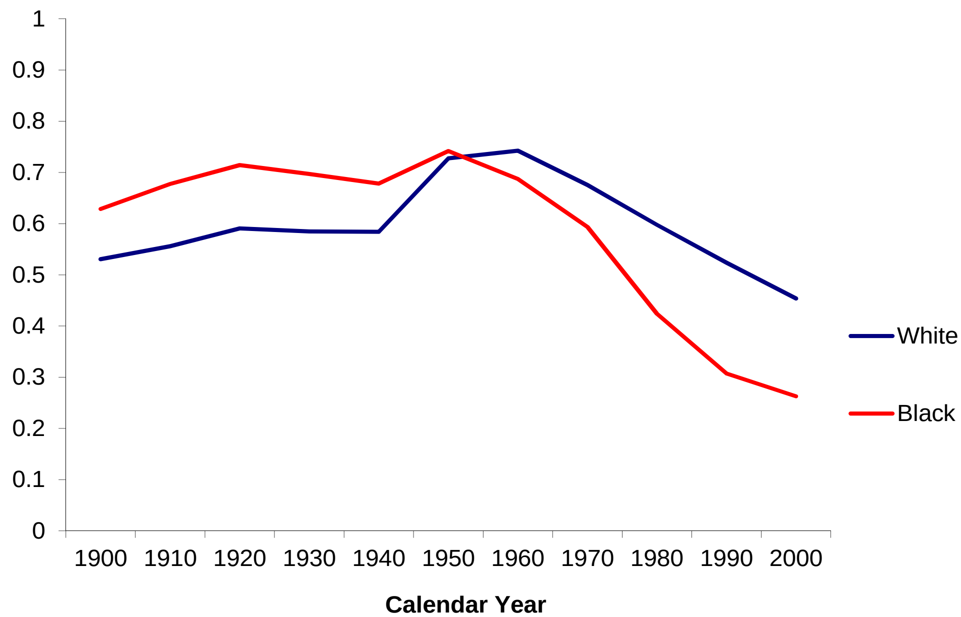
Parameters	Estimates	Standard Errors
First Shape Parameter	0.020	[0.011]
Second Shape Parameter	0.072	[0.044]
Value of Being Single	0.107	[0.039]
Fraction of Men Unwilling to Marry	13.3	[0.159]

Figure 1: Number of Marriages



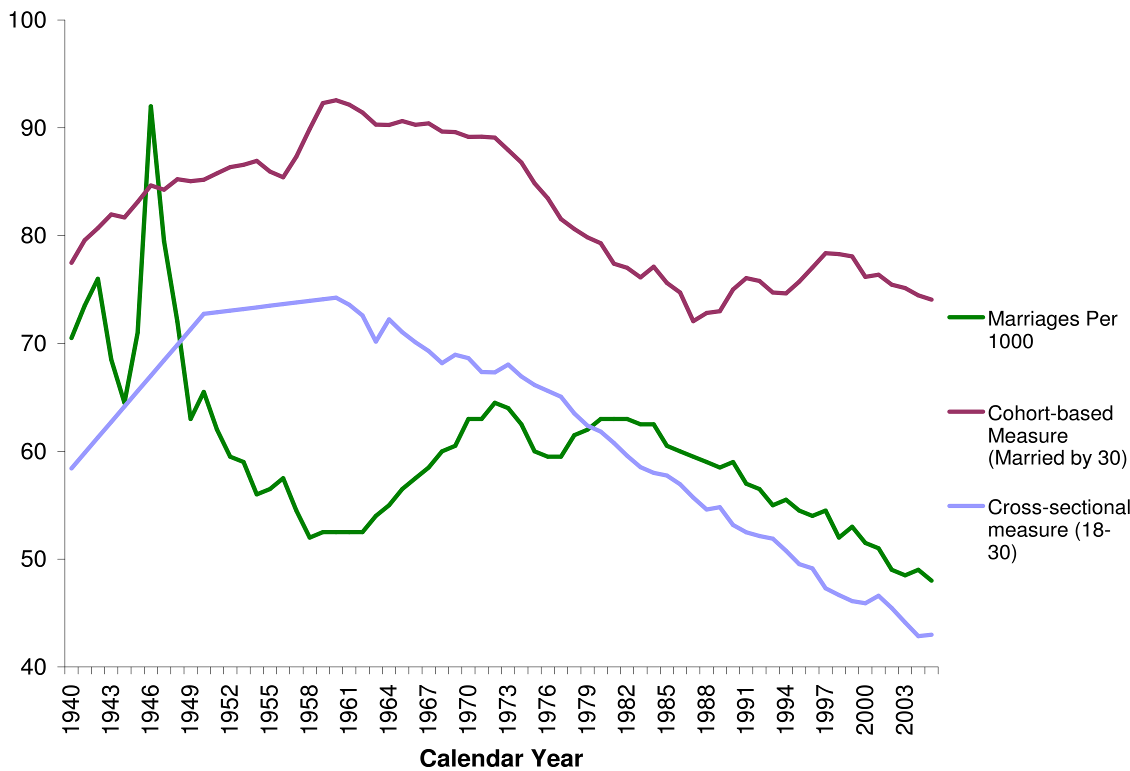
Source: Vital Statistics of the United States.

Figure 2: Share of Women Ever Married, Ages 18 to 30



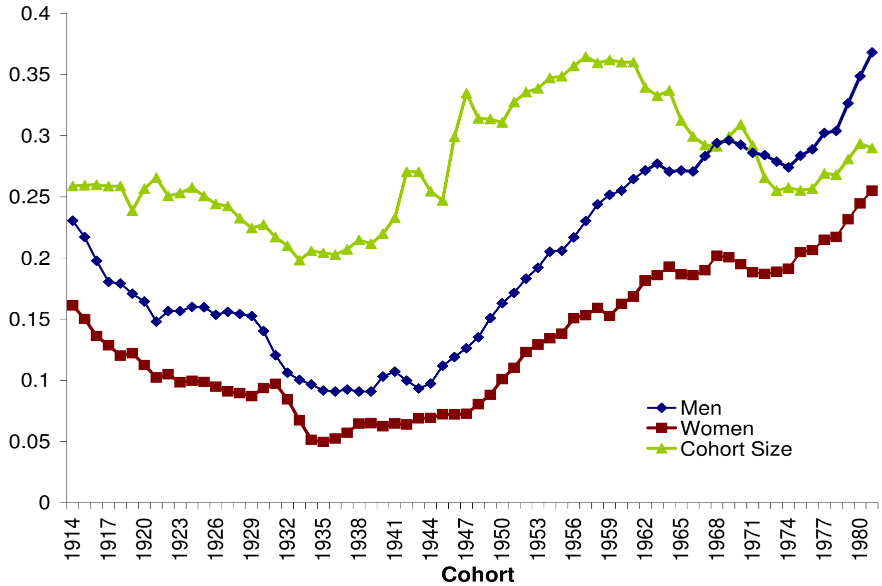
Source: IPUMS USA Census, 1900-2000.

Figure 3: Different Measures of Marriage Rates

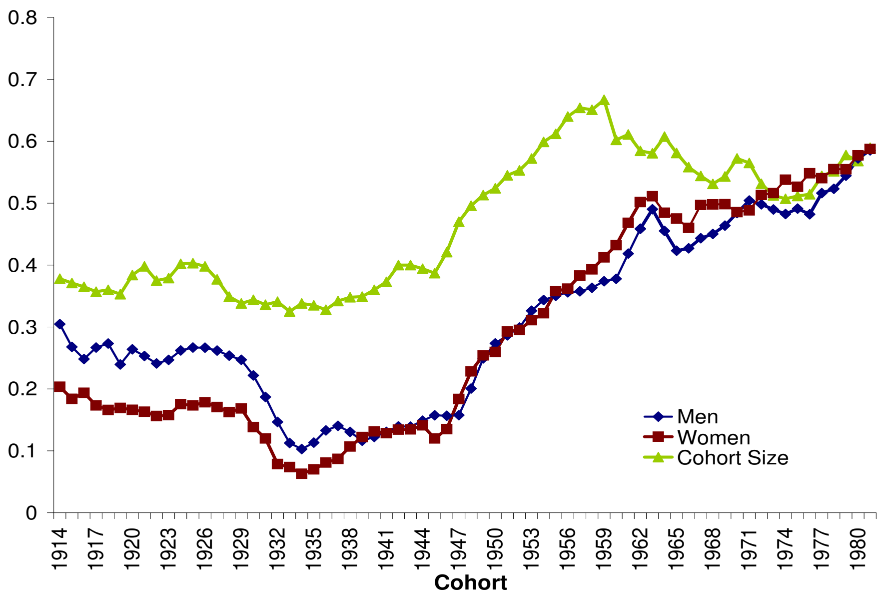


Sources: Vital Statistics of the United States; IPUMS CPS, 1962-2005; IPUMS USA Census, 1940-1960. The vertical axis corresponds both to the marriages per thousand, as well as to the percentage of individuals ever married for the cohort-based and cross-sectional measures. To facilitate a direct comparison between the three measures, cohabitations are not accounted for in any of the measures.

Figure 4: Share Never Married By 30



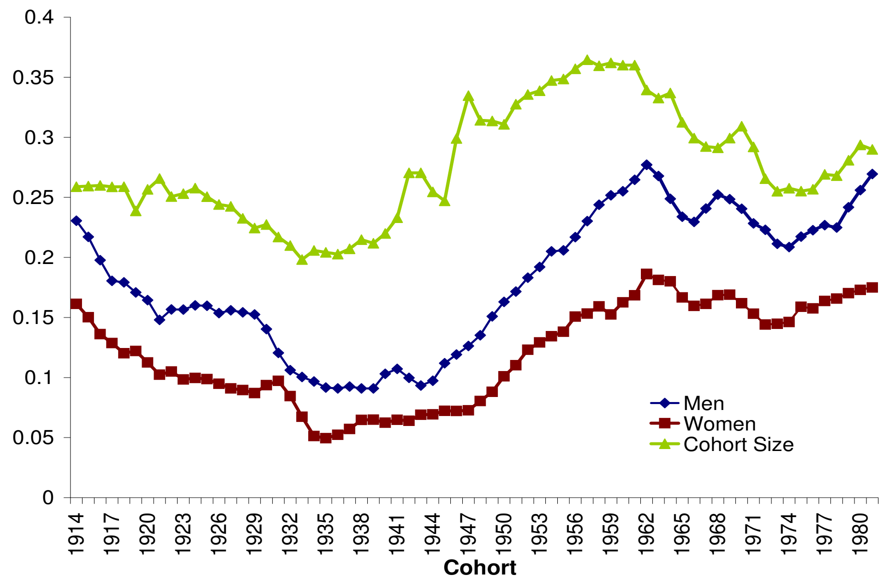
(a) White



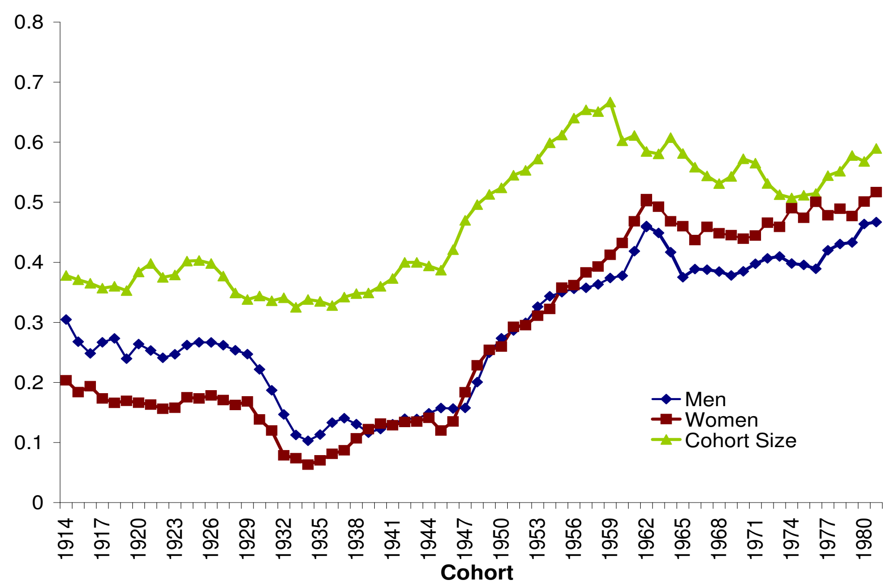
(b) Black

Sources: Vital Statistics of the United States; IPUMS CPS, 1962-2011; IPUMS USA Census, 1960-1980.

Figure 5: Share Never Married and Not Cohabiting By 30



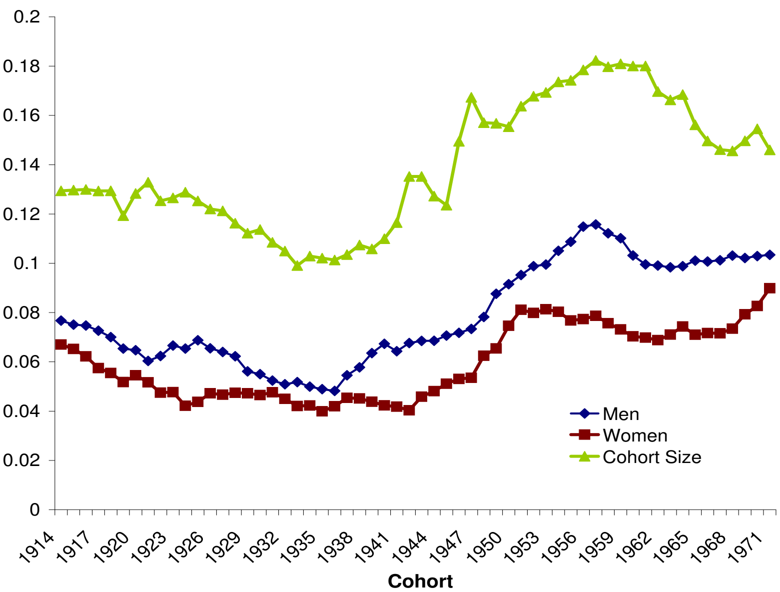
(a) White



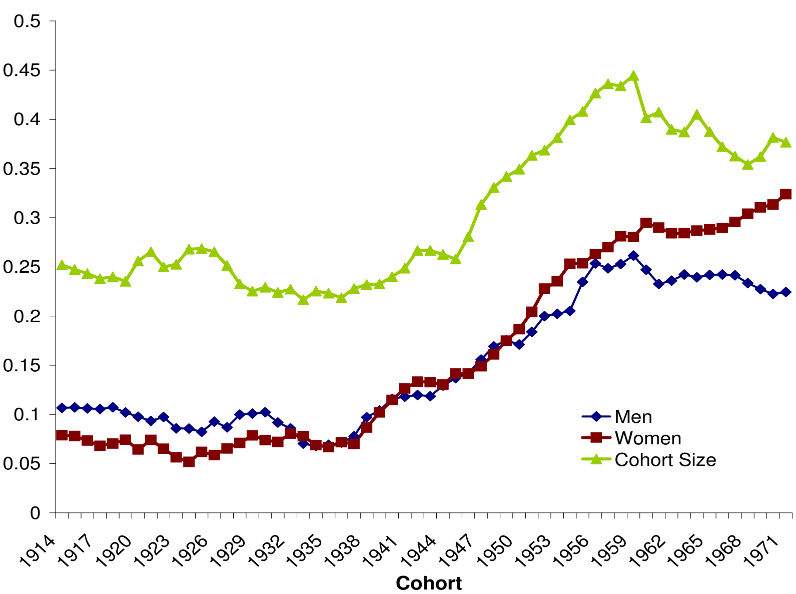
(b) Black

Sources: Vital Statistics of the United States; IPUMS CPS, 1962-2011; IPUMS USA Census, 1960-1980.

Figure 6: Share Never Married and Not Cohabiting By 40



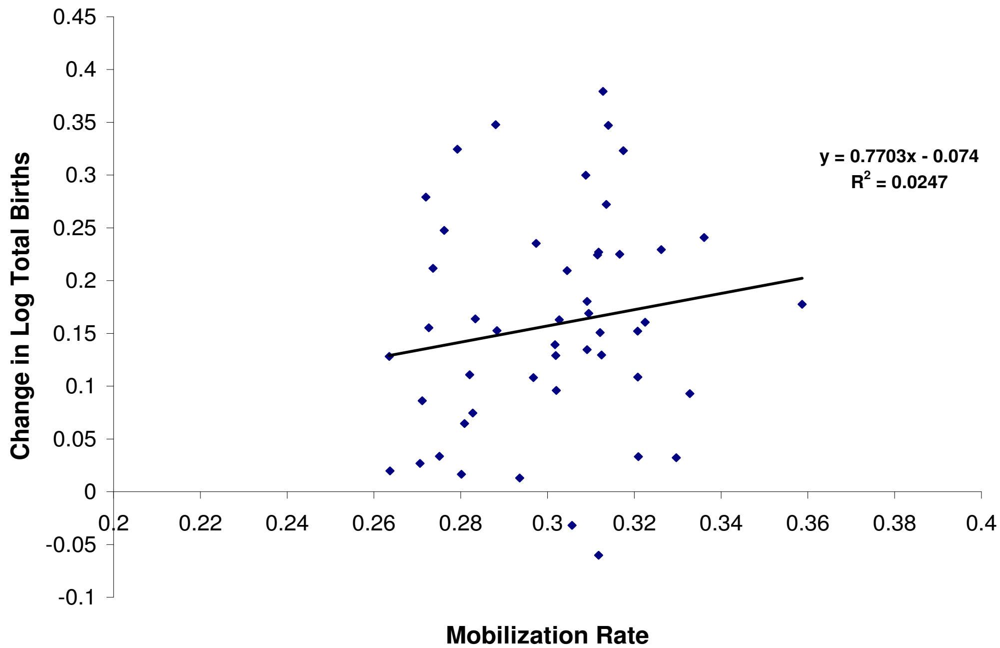
(a) White



(b) Black

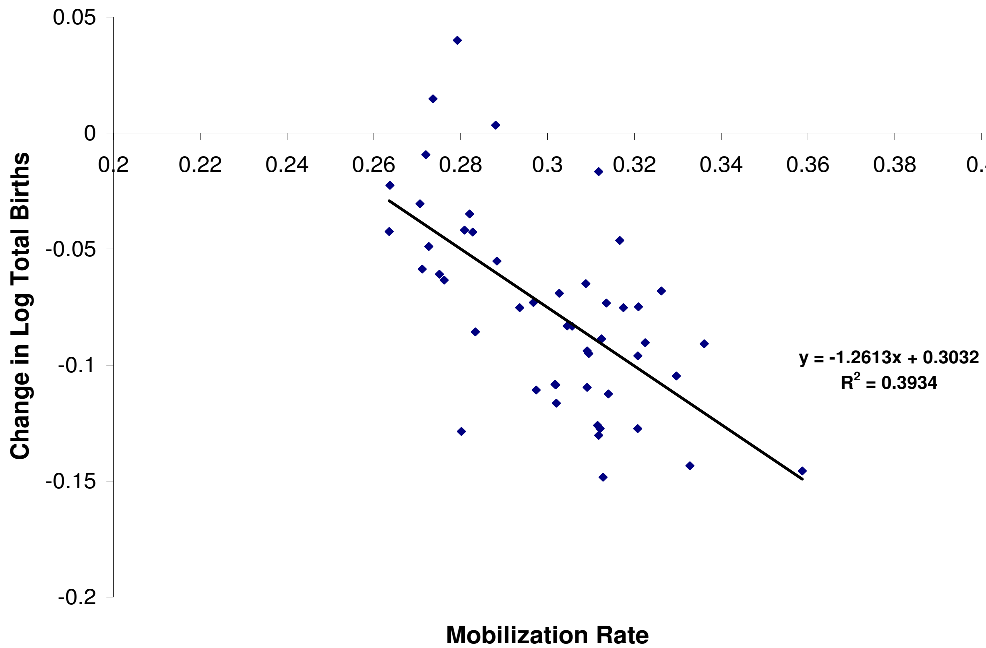
Sources: Vital Statistics of the United States; IPUMS CPS, 1962-2011; IPUMS USA Census, 1960-1970.

Figure 7: Mobilization Rates and Changes in Total Births, 1940-1942



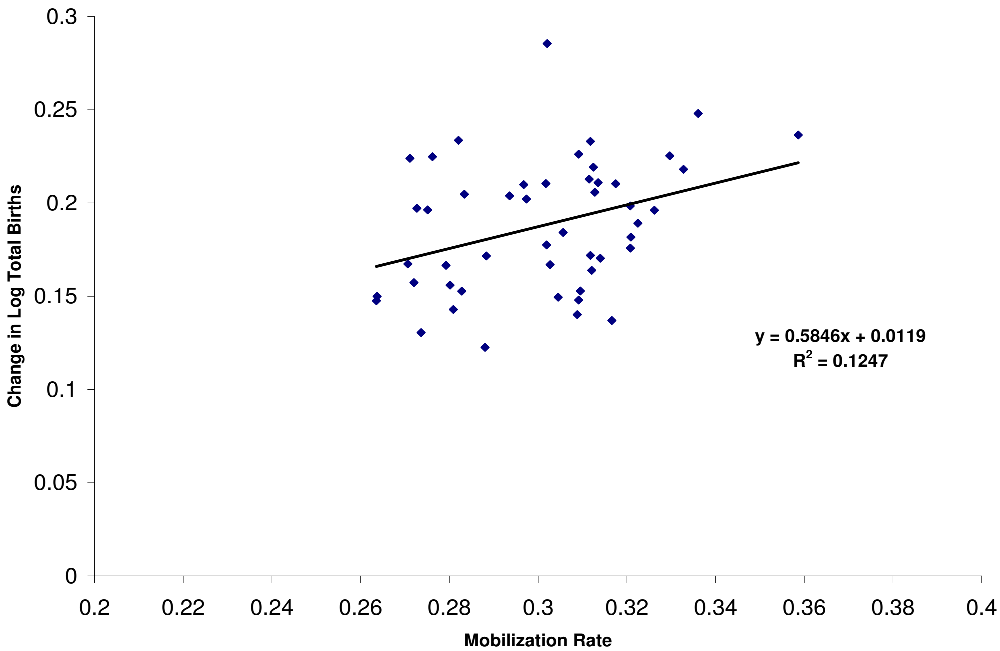
Source: Vital Statistics of the United States; Special Monograph of the U.S. Selective Service System, 1956. Births and induction rates are for whites only.

Figure 8: Mobilization Rates and Changes in Total Births, 1943-1945



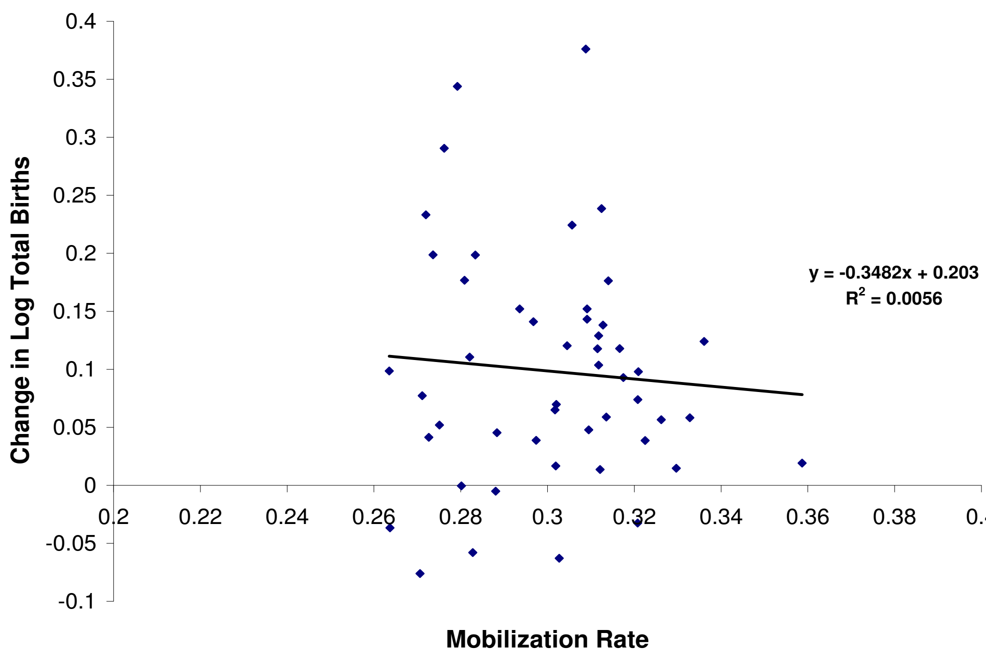
Source: Vital Statistics of the United States; Special Monograph of the U.S. Selective Service System, 1956. Births and induction rates are for whites only.

Figure 9: Mobilization Rates and Changes in Total Births, 1945-1946



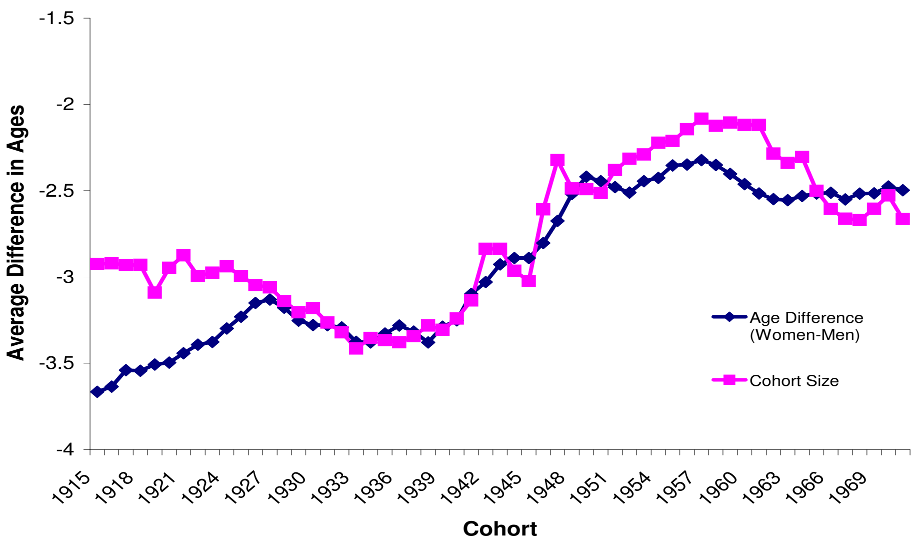
Source: Vital Statistics of the United States; Special Monograph of the U.S. Selective Service System, 1956. Births and induction rates are for whites only.

Figure 10: Mobilization Rates and Changes in Total Births, 1935-1940



Source: Vital Statistics of the United States; Special Monograph of the U.S. Selective Service System, 1956. Births and induction rates are for whites only.

Figure 11: Age Difference Between Spouses by Cohort, Whites



Source: Vital Statistics of the United States; IPUMS CPS, 1962-2011; IPUMS USA Census, 1960-1980.

Figure 12: Observed and Simulated Never Married Rates