# Cohort Size and the Marriage Market:

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

Mary Ann Bronson† Maurizio Mazzocco‡

Current Draft, September, 2013. First Draft, February, 2012

**Abstract**

We propose an explanation for almost a century of changes in U.S. marriage rates, in three stages. First, we show that changes in cohort size alone can account for around 50 to 70% of the variation in marriage rates since the 1930s for both black and white populations. Specifically, increases in cohort size reduce marriage rates, whereas declines in cohort size have the opposite effect. We provide the most convincing evidence on this relationship by using variation in cohort size due to differences across states in sale bans on oral contraceptives. Using this exogenous variation in access to oral contraceptives, and consequently the number of births, we provide evidence that the relationship between changes in cohort size and changes in marriage rates is causal. Next, we develop a dynamic search model of the marriage market that qualitatively generates this observed relationship, and derive a testable implication about cohort size's effect on spouses' age differences. Finally, we estimate the model and investigate its consistency with the data. We fail to reject it using the derived implication, and find that it can quantitatively explain much of the observed variation in marriage rates.

## 1 Introduction

What causes variation in marriage rates over time? For both economists and policy-makers, this is a question of significant interest, as a large body of evidence suggests that marriage rates have important implications for other economic variables. Such variables include fertility rates, children's welfare, children's education, labor force participation, hours of work, income inequality, the fraction

*We are grateful to Ran Abramitzky, Guillermo Beylis, Pierre-André Chiappori, Edgar Cortes, James Heckman, Joe Hotz, Caroline Hoxby, Kenneth Mirkin, Anne Pebley, Luigi Pistaferri, Dmitry Plotnikov, Nico Voigtlaender, and participants at various seminars and conferences 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.

of individuals on government aid, population growth, and workers' productivity.[1] In spite of this, according to the existing literature, no prevailing explanation can account for the variation in marriage formation over time and across geographies. Many existing theories about changes in the marriage rate, which will be reviewed in the next section, have either been empirically rejected, or have explanatory power that is limited to specific periods or specific groups of individuals.

The main contribution of this paper is to provide an explanation for changes in U.S. marriage rates that holds empirically over nearly a century. We show that one variable, changes in cohort size, explains the majority of the variation in U.S. marriage rates since the early twentieth century, both over time and across states. 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 two steps. First, using both time-series variation and cross-sectional state-level variation in cohort size and marriage rates we find that there is a strong and negative relationship between marriage rates and cohort size. On average, a 10-percent increase in cohort size is associated with a 2 to 3 percent decrease in the share ever married by 30 and around a 0.5 to 1 percent decrease in the share ever married by 40. These are sizeable effects that account for around 50 to 70 percent of the variation in marriage rates since the early twentieth century. Our results indicate that changes in cohort size account well for medium- and long-term changes in marriage rates, but cannot explain the short-run year-on-year variation. They also suggest that part of the effect of cohort size on marriage rates operates through an increase in the age at first marriage.

In the second 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 based on Bailey (2010), we employ the interaction between the 1957 introduction of *Enovid*, later known as the birth-control pill, and cross-state variation in anti-obscenity laws, which limited the use of contraception in some states until the mid-1960s, to generate exogenous variation in the number of births and therefore cohort size. Our results indicate that, in states that limited contraceptives, cohort size increased relative to states that did not have such limits, and that this change generated a decline in marriage rates. 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 justified by the fact that men's fertile lifespan is longer than women's, and the common insight that one benefit of marriage

[1] See for example Killingsworth and Heckman (1986), Moffitt (2000), Angrist and Lavy (1996), Gruber (2004), McLanahan and Percheski (2008).

is that it is an effective arrangement for raising children. 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 intuition underlying this result is straightforward. In the model, the marriage market is populated by young women and by young and old men. A positive shock to cohort size with some degree of persistence will therefore have two effects. First, old men become scarce relative to young women. As a consequence, young women will be less likely to meet and marry them. This direct effect will reduce the marriage rate. Second, young men will become more selective. The reason for this is that they will have a higher probability of meeting a young woman in case they choose to wait until they are old. As a consequence, if they meet a potential spouse whose quality is not sufficiently high, they will decide to wait. This indirect effect will further reduce the marriage rate. The model can therefore qualitatively explain the negative relationship between cohort size and marriage rates. We then derive an implication that can be used to test the model. We show that in our 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 we do not reject it. 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 a substantial portion of 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.

Our findings have an important policy implication. In recent years, politicians and policy makers have started to discuss and implement policies that attempt to improve the well-being of low income families by increasing the fraction of married individuals. For instance, during the Bush Administration, proposals pending approval planned to allocate up to 1.5 billion dollars to implement and evaluate policies aimed at promoting marriage.[2] Our results suggest that these policies may prove to be largely ineffective since a significant part of the changes in marriage rates is generated by forces that are mostly outside the control of policy makers. This is likely to be particularly true in the short and medium term which is the time horizon available to politicians and policy makers.

We conclude this section with one remark. The variable cohort size is not exogenous. It is affected by variables like new fertility technologies, technological progress, child care availability, supply of housing, and changes in labor supply decisions of women. This fact does not diminish the importance of our results for social scientists and policy makers. They indicate that to understand the dynamics

[2] Seefeld and Smock (2004) provide a nice discussion of the recent interest of policy makers in marriage as a policy tool.

of the marriage market one has to study the evolution of cohort size and of the variables that have an effect on it.

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 some of the existing explanations for the variation in U.S. marriage rates for which some empirical evidence is provided.

One set of explanations for historical changes in marriage rates focuses on changes in income. Cherlin (1981) among others notes the correspondence between rising incomes after World War II and the associated marriage boom during this period. A related common insight is that low 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. These results suggest that income may generate part of the variation in marriage rates. But they also suggest that variation in income cannot be the general explanation that underlies the changes in marriage rates observed in the past century. In addition, there is a potential reverse causality problem with this theory that is not addressed: men have higher labor hours and earn more following marriage.[3] It is therefore difficult to determine whether an increase in income causes marriage rates to rise or whether an increase in the marriage rate generates higher income levels.

An alternative set of explanations focuses on technological innovations. Akerlof, Yellen, and Katz (1996) consider the adoption of new fertility technologies such as the pill and abortion in the sixties and seventies. The authors argue that these technologies increased the number of out-of-wedlock births and reduced the number of marriages. A potential direct effect on marriages could be that the adoption of the new technologies mechanically reduced conceptions and therefore 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 decrease the incentives of women to obtain a promise of marriage if premarital sexual activity results in pregnancy with a consequent rise in out-of-wedlock

[3] A number of studies have documented this relationship. See, for example, Mazzocco, Ruiz, and Yamaguichi (2007).

births and decline in marriage rates. This explanation may provide important insight into changes during and around the 1970s, but it would not be able to explain intervals of increasing marriage rates after the 1970s, and it cannot address historical variation prior to this period.

Greenwood and co-authors also focus on technological innovations and suggest that the decline in the price of household appliances explains patterns for several household outcomes, including marriage. Greenwood and Guner (2009) argue that labor-saving technological progress in the household sector made it easier for singles to maintain their own home, increasing the value of being single and reducing the marriage rate. Their theory addresses in particular the period of rapidly falling marriage rates starting around the mid-sixties. However, the theory would have trouble explaining much of the remaining historical variation in marriage rates. In as far as technological progress has constantly improved household appliances in the past 100 years, we might predict that marriage rates either decline systematically throughout this period or decline and then remain constant. However, this contradicts what we observe in the data.

A different well-known set of theories considers men's marriageability and policies such as welfare aid, both of which may affect women's desire to marry. These theories have been widely empirically tested, with mixed results. Ellwood and Crane (1990) review papers that have tested the hypothesis proposed by Wilson (1987) that limited labor market opportunities reduce the number of marriageable men, and thus the marriage rate. While some papers provide evidence in support of this theory, others reject this hypothesis (e.g., Plotnick (1990) and Lerman (1989)). The degree to which men's employment opportunities affect marriage rates is largely still an open question. A related issue affecting especially black men's marriageability in the last three decades is the rise in incarceration. Charles and Luoh (2010) study the relationship between incarceration and marriage rates and find that higher incarceration rates decrease the fraction of married individuals both in black and white populations. Quantitatively, however, such an explanation is primarily relevant for black populations and only for the period after 1980, when drug-related policies significantly increased incarceration rates. Ellwood and Crane (1990) have also evaluated the link between welfare aid and marriage. They conclude that there is very little empirical support for the proposition that welfare benefits played a major role in marriage trends for black or white women. Finally, rising income and employment opportunities for women may also affect their desire to marry. However, we do not know of papers that formally test this hypothesis over time or across geographies. One reason may be that potential reverse causality complicates such an empirical analysis: women who face poorer marriage prospects may both invest more in human capital and work more.

Finally, there are two explanations that have 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 many of the variables that determine the economic and social outcomes of a birth cohort:

earnings and unemployment rates, college enrollment rates, divorce, fertility, crime, suicide rates, and marriage. Easterlin's explanation is that when income is above the aspiration level for a given cohort, the individuals in that 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 outcomes. 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 (Pampel and Peters (1995)).

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 marriage 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. In a well-known series of case studies, Guttentag and Secord (1983) argue that declines in the number of women relative to men lead to higher marriage rates, among other social changes. 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. 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 marriages were often formed between individuals belonging to the same ethnicity, and finds that ethnicities that experienced lower sex ratios display higher marriage rates. Finally, Abramitzky, Delavande, and Vasconcelos (2011) consider variation in the sex-ratio due to World War I casualties in France and find that a higher sex ratio is associated with a lower marriage rate for women and with a larger marriage rate for men. In the search model we develop, variations in marriage rates are determined by changes in the sex ratio, which are endogeneosly generated by changes in cohort size. The mechanism we propose is therefore partially related to papers that use changes in the sex ratio to explain changes in marriage rates.[4]

[4]Several papers have studied the relationship between the sex ratio and other economic variables. One example is the paper by Neal (2004) where the author studies the relationship between sex ratios and rates of single motherhood. A second example is the paper by Grossbard-Shechtman (1984) where the author analyzes the link between the sex ratio and labor force participation of women. Another example is the paper by Bitler and Schmidt (2011) which shows that 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.

## 3 Data

In this section we describe the data used in the analysis. Throughout the paper we rely on the following datasets: National Vital Statistics (1909-1980), Census total population counts (1910-1980), Survey of Epidemiology and End Results (SEER) population estimates (1969-2000), IPUMS CPS (1962-2011) and IPUMS USA (1940-2000). In the Data Appendix B we provide a detailed description of how the datasets are used to construct the main variables employed in the empirical analysis. In this section, we give a brief summary of that description.

In the empirical analysis we employ two main variables: cohort size and the share ever married by a given age for a given cohort. We construct two different measures of cohort size: cohort size at birth and cohort size at marriageable age. The first variable is used with longitudinal variation, whereas the second one is employed with cross-state variation. In this paper we are interested in the evolution of the variable cohort size at the ages in which individuals choose whether and whom to marry. With longitudinal variation, however, we use cohort size at birth as the main independent variable for two reasons. First, as shown in Figure 1, when cohort size is computed for the U.S. population there is little difference between cohort size at birth and cohort size at marriageable age, since migration from and to the U.S. was limited during the time period we consider. Second, we can construct the variable cohort size at birth for cohorts born in 1909 and after. The variable cohort size at marriageable age can only be constructed for cohorts born in 1940 and after. By using cohort size at birth we can therefore consider a larger number of cohorts without significant effect on the analysis. When we use cross-state variation, we have to use the variable cohort size at marriageable age because of large migration flows across states during the period of investigation.

The variable share ever married at age 30, 35, and 40 is constructed using either the decennial Censuses or the SEER population estimates. Appendix B describes the exact procedure used to construct this variable.

## 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 compare them to our alternative measure, which we believe is better suited to the examination of marriage choices over time. We then provide empirical evidence on the relationship between cohort size and marriage rates using longitudinal variation. In the third subsection, we describe 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 by using variation in early access to oral contraceptives across states as a plausible source of exogenous variation in cohort size.

## 4.1 Different Measures of Changes in the Marriage Rate

When analyzing changes in marriage rates over time, existing studies typically employ one of the following two variables: the number of new marriages per population; and the share of individuals currently or ever married within some age range, e.g. between the ages of 18 and 30. We use a different measure, the share of individuals in a given cohort ever married by a given age. In this section we compare the behavior of these three measures. As a first contribution, we show that, when used to study the evolution of the marriage market over time, the commonly used measures are problematic because they confound different effects and can potentially lead to incorrect inference about overall marriage behaviors.

Figure 2 graphs the three measures described above: the number of marriages per 1,000 individuals; the share ever married in a cross-section of women between the ages of 18 to 30; and the share of women in a given cohort ever married by 30. We refer to these measures as the population-based measure, the cross-sectional measure, and the cohort-based measure, respectively. The first two measures are based on the calendar-year, whereas our measure is based on the cohort's year of birth. To make them comparable, we add 25 years to the cohort's birth year and plot our measure together with the other two. It is immediately evident from Figure 2 that the three measures give very different pictures of how marriage rates in the U.S. have evolved over time. It is worthwhile to discuss why they behave differently and the main problems with each measure.

The first measure, the number of marriages per 1,000 individuals, is based on the number of new marriages that take place in a given year. It is available from the U.S. Vital Statistics. This measure can be useful to detect years that experience an unusual growth or decline in marriage rates. For example, it captures well the rapid rise in the number of marriages in the immediate post-war years. However, this measure is not as well suited to study the evolution of the share ever married for two main reasons. First, using the population-based measure it is impossible to distinguish between first, second, or later marriages. This distinction is important if the researcher is interested in evaluating the fraction of people who choose to marry, since the second and later marriages should not be included in a measure designed to describe the evolution of the share ever married. The second and more serious problem for evaluating marriage trends is that this measure conflates changes in the numerator, new marriages, with changes in the denominator, population. As a result, if the population undergoes any substantive growth or decline due, for instance, to changes in fertility or migration patterns, one may draw the wrong inference. For example, starting from 1946 this variable drops steeply for about fifteen years. One might infer that marriage rates were falling steadily in the forties and fifties, but both the cross-sectional measure and the cohort-based measure show that marriage rates were flat or rising during this time. Part of the explanation for the large decline in the population-based measure 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 population-based variable displays rapid growth which can been interpreted as a big increase in marriage rates. The cross-sectional and cohort-based measure, however, show that this interpretation is misleading. During this period, marriage rates experienced a slight decline. A potential explanation for the rise in the population-based measure is that the U.S. population declined during the baby bust that characterized the U.S. in the sixties.[5]

The second measure illustrated in Figure 2 is the share ever married in a cross-section of women between the ages of 18 to 30, which was constructed using the CPS and Census. 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 cross-sectional measure 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 the divergence is that the cross-sectional measure is strongly affected by changes in the age at first marriage: when examining the share of people ever married within an age range, one cannot determine whether individuals are simply delaying marriage or whether they choose not to marry. Starting from the second half of the eighties, the age at first marriage experienced a significant increase, which explains why the cross-sectional measure declines during this period whereas the cohort-based measure increased.

It is important to note that the cohort-based measure may also be affected by changes in the age at first marriage if the age cut-off is too low. It is therefore important that one chooses properly the age cut-off: if one believes that an age cut-off of 30 is too low, an age cut-off of 35 or 40 should be used. In the next subsections, we will show that the trends the cohort-based measure captures look very similar at different age cut-offs. In Figure 2, we choose an age cut-off of 30 mostly for expositional purposes, so that we can include as many recent cohorts as possible.

Because of the advantages the cohort-based measure has over the other two variables, throughout the rest of the paper we will only use that measure to study the evolution of marriage rates.

## 4.2 Change in Marriage Rates Over Time

In this subsection we provide evidence on the relationship between changes in cohort size and changes in marriage rates using longitudinal variation. The evidence is presented in two steps. We first provide evidence on the general nature of this relationship. We then try to understand whether cohort size can explain the short-run, medium-run, or long-run changes in marriage rates.

In Figure 3 we plot cohort size and the share never married by age 30, separately for women and men, for all cohorts born between 1914 and 1981. The first panel describes these variables for the

[5]A second potential explanation for the increase in the population-based measure in the seventies is that the large baby boom cohorts were coming of age. Even though these individuals were marrying at lower rates, their large overall number will generate and increase in the number of registered marriages, the numerator.

white population whereas the second panel plots them for the black population. We plot the share *never* married because visually it is easier to detect a positive correlation between the two variables. Figure 3 contains one main finding. To describe it, we initially focus on cohorts born before 1960. For those cohorts, there is a strong positive correlation between cohort size and the share never married. The decline in size for cohorts born in the 1920s and 1930s is associated with a similar drop in the share never married. This decline corresponds to the well-documented "marriage boom" that starts in the mid-1940s and lasts through the early 1960s, the period in which the cohorts born in the twenties and thirties were active in the marriage market. The sharp increase in the size of cohorts born between 1946 and 1959, which correspond to the post-war baby boom generations, is associated with a share never married that nearly tripled during this period. Births and the share never married for the black population follow similar patterns.

It is left to explain why we lose the positive correlation between our two main variables for cohorts born in the 1960s and 1970s. Note that those cohorts were active in the marriage market starting from the 1980s, which is the period in which cohabitation began to become a popular form of household formation and potentially a close substitute for marriage. To understand whether cohabitation can resolve the inconsistency between the early and later cohorts, in Figure 4 we plot the variables reported in the previous figure, with the exception that now cohabiting individuals are treated as married individuals instead of being treated as never-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 early 1970s correspond to a decline in the share never married and not currently cohabiting by 30. Increasing cohort sizes in the second part of the 1970s are associated with a rise in the share never married and not cohabiting.

Figure 4 contains a second noteworthy finding. The strong positive 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 markets because it challenges the common perception that the two markets are governed by different rules and exhibit different marriage behaviors. Figure 4 suggests that the two marriage markets are similar in at least one respect: they respond to changes in cohort size in a similar way.

In Figures 3 and 4, 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 5 we plot cohort size and shares never married and not cohabiting by age 40. Using this new cutoff age, we find patterns that are similar to the ones 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 for age cutoffs of 30, 35, and 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 30, 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, where an increase in cohort size is 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. An increase of 10% in cohort size reduces the share ever married or cohabiting by 40 by a percentage that ranges from 0.66% for white women to 4.53% for black women. In percentage points, this amounts to a decline that is between 0.6 and 3.8 points 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.[6]

In the rest of the paper, we will focus on the white population only for one main reason. White women and men have similar cohort size at the time of marriage, whereas black men have a significantly lower cohort size than black women because of higher mortality and incarceration rates. As a consequence, the investigation of the marriage market for blacks requires a different type of analysis which we undertake in a separate paper.

In the remaining part of the section we will try to understand whether cohort size can explain the short-run, medium-run, or long run changes in marriage rates. Observe that the regressions in log-levels documented in Table 1 capture the short-run, the medium-run, as well as the long-run effect of changes in cohort size on changes in marriage rates. To see this, note that the regressions in levels measure the correlation between changes in our two main variables independently of when the changes took place. The same weight is assigned to the change in these two variables between the cohort born in 1935 and the cohort born in 1958 as to the change between the cohort born in 1935 and the cohort born in 1936. To measure the short-run, medium-run, and long-run effects of changes in cohort size on changes in marriage rates, we regress $n$-year differences in marriage rates on $n$-year differences in cohort size where $n$ is set equal to 1, 2, 3, 4, 5, 7, and 10. To capture the effect of adjacent cohorts, for $n > 1$ we use differences in cumulative cohort size as our independent variable, where cumulative

[6]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.

size for cohort born in period $t$ for the $n$-year difference is constructed by adding up cohort size from $t-n+1$ to $t$.[7] We interpret the coefficient estimates on the 1-year and 2-year differences as the short-run effect of cohort size on marriage rates, the coefficients on the 3-year, 4-year, and 5-year differences as the medium-term effect, and the coefficients on the 7-year and 10-year differences as the long-term effect. This choice is somewhat arbitrary, but it helps us focus the discussion.

The results are presented in Table 2. The first two columns report the short-run effect. Clearly, in the short run cohort size has at best a weak effect on marriage rates. The estimates for the 1-year differences indicate that a 1-year change in cohort size is not sufficient to trigger a change in marriage rates. With the exception of the coefficient for women by age 30 and men by age 40, all the coefficients for the 1-year difference are statistically equal to zero. The estimated coefficient for men by age 40 is the only one in all of the estimations we have performed that is positive and statistically significant. This result is generated by the two spikes in births that occurred in 1942 just at the start of World War II and in 1946-1947 after World War II ended. If one drops the observations that characterize the period around World War II, the coefficient becomes zero. The effects are slightly larger when we employ 2-year differences. Now the coefficient for women by age 35 is also negative and statistically significant. But even for 2-year differences, the effect of cohort size is very weak. The effect of cohort size on marriage rates is much stronger when we study the medium-term effect. With 3-year differences all the coefficients become large in size, negative, and statistically significant. The only exceptions are the coefficients by age 40 which are negative but statistically not significant. When we increase the differences to four and five years, the coefficients become larger in size and are now all statistically significant. The long-term effect of cohort size is even stronger. The coefficient estimates for the 7-year and 10-year differences suggest that, for an age cutoff of 40, an increase in cohort size of 10% generates a drop in marriage rates of 0.6-0.8%. This is a significant decline since it implies that a standard deviation increase in cohort size decreases the share ever married by 40 by a third to a half of a standard deviation. These findings indicate that cohort size can explain the medium and long term variation in marriage rates, but not the short term variation. Changes have to cumulate for longer than one or two years to generate significant fluctuations in marriage rates.

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 medium-run and long-run time series variation in marriage rates. In the following subsections, we further explore the empirical link between cohort size and marriage rates. Because cohabitation has become a close substitute for marriage since the early eighties, in 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

[7] We have also estimated the effect of cohort size using simple $n$-year differences. The estimates display similar patterns, but the coefficients are generally smaller and are less precisely estimated.

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 larger increases in cohort sizes experience larger drops in marriage rates.

To use cross-state variation, there is one issue we need to address that is not present in the longitudinal analysis. Changes in the size of a cohort at the time individuals are of marriageable age are endogenous because they are partially driven by migration decisions. 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. To address this potential source of endogeneity, we use total births in a given year and state, which are arguably unaffected by the endogeneity issues discussed above, as an instrument for the size of the marriage market.

To perform the analysis at the state level, we rely on the decennial Census over the entire period of interest since it is the only dataset with sufficiently large sample sizes for all states. Appendix B provides details about how we construct the three main variables needed for the analysis: cohort size at birth, cohort size at marriageable age, and share ever married for each state and cohort. Using the Census data we can only use the decennial cohorts born between 1910 and 1970, since the share ever married can only be computed for them. We therefore perform the empirical analysis by first constructing ten-year differences for the log of share ever married, the log of cohort size at birth, and the log of the size of the marriage market for each cohort and state. We then use these variables in our regressions. Because of the small number of observations, we pool all cross-sections and regress the differences in log share ever married on the differences in log cohort size where log cohort size is instrumented using log cohort size at birth. We add year fixed effects to the regression to control for general time trends.

Table 3 presents the results of the cross-state regressions separately by gender. The first column presents the estimates obtained using a standard OLS regression of 10-years differences in log share ever married on 10-year differences on the size of the marriage market. Similarly to the findings obtained using longitudinal variation, the estimated coefficient is negative and statistically significant for the two age cutoffs we consider, for both men and women. As argued above, the OLS estimates may be biased because of migration decisions. In particular, if migration is partially driven by the desire to find a spouse, the estimates will be downward biased. A similar result applies if migration is partially driven by job-related reasons. In this case, on average one would expect individuals to move to states with higher earning opportunities. If individuals with higher income are also more likely to marry, this would generate a positive correlation between cohort size growth and marriage rates, biasing the results away from the strongly negative relationship we predict.

We can now describe the estimates obtained when we control for endogeneity by instrumenting the size of the marriage market using cohort size at birth. At the bottom of the table we present the first stage results. Two findings are worth discussing. First, the first stage suggests that cohort size at birth explains a large fraction of the marriage market size. Second, the coefficient on log cohort size at birth is estimated to be 0.44, which is well below 1. This indicates that cross-state migration plays an important role in changes in cohort size at marriage age. In column two we present our IV estimates. They are all negative, statistically significant and, as expected, larger in size than the OLS estimates. For the share ever married by 30, the coefficient is estimated to be -0.104 for women and -0.123 for men. When we use the age cutoff of 40, the coefficient drops to -0.058 for women and -0.047 for men. The size of the IV estimates are therefore similar to the ones obtained using the longitudinal data. They also similarly decline when we increase the age cutoff. In the third column, we add time-region fixed effects to the IV specification to make sure that our findings are not driven by systematic differences in trends across regions. The coefficient estimates are similar to the ones presented in column 2, but we lose significance for men when we use an age cutoff of 40 because of an increase in the standard errors.

### 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 address the potential endogeneity issues outlined in the previous subsection, we construct an instrument which is based on an idea first proposed by Bailey (2010). The idea is to use the interaction between the introduction of *Enovid* in 1957, later known as the the birth control pill, and cross-state variation in anti-obscenity laws, which limited the use of contraception, to generate exogenous variation in number of births and therefore cohort size.

In 1873, the U.S. Congress enacted the Comstock Act which had two main goals. The direct objective was to ban the interstate mailing, shipping, and importation of products and printed materials that were considered to be "obscenities". Since the Act considered anything employed for the prevention of conception an obscenity, it outlawed any interstate transaction involving contraceptives. The indirect objective was to "incite every State Legislature to enact similar laws" as stated by U.S. Representative John Merriman during an interview with the New York Times on March 15, 1873. The Comstock Act was highly successful in achieving this goal. By 1900, 42 states had approved anti-obscenity laws and by 1943 the number of states had increased to 48.

There is considerable variation across states in the type of anti-obscenity statutes that were enacted. As a consequence, these laws had different effects on the introduction of the pill in different states. As suggested by Bailey (2010), the states can be grouped into four categories depending on the type of law they enacted. The first group includes states that explicitly banned the sale, advertisement, and distribution of information of any product for the prevention of conception. This category includes seventeen states. The second group consists of all states that had the same ban on sales, advertisement, and distribution of information as the first group of states, but added an exception for physicians and pharmacists who were allowed to sell, advertise, and distribute information on products and materials related to birth-control methods. Seven states belong to this category. The third category includes states that explicitly only banned the advertisement and distribution of information of products and materials for the prevention of contraception, but did not outlaw their sale. Six states enacted this type of statute. The final group is composed of states that approved a law that banned the sale, advertisement, and dissemination of information of obscene products and materials, without explicitly classifying the prevention of conception as obscene. This category includes eighteen states. In our analysis, we refer to states in the first two groups as having a sales ban. We control explicitly for whether or not a state had a physician exception.

An important question is whether some states enacted stricter anti-obscenity laws because they had more conservative constituencies or because of other observable cross-state differences. Bailey (2010) provides evidence that this is not the case. For instance, among the states that adopted sales bans of contraceptives one can find both typically conservative and typically liberal states. California and Washington, two of the states that repealed anti-abortion laws before the Roe v. Wade decision, enacted the strictest version of the bans whereas Alabama, a generally conservative state, adopted a statute that did not explicitly categorize the prevention of conception as obscene.

These anti-obscenity state laws lasted until the sixties when they were repealed or struck down by the individual states or by the 1965 U.S. Supreme Court's decision in *Griswold v. Connecticut*. Specifically, two states repealed their anti-obscenity statutes in 1961, one state in 1962, four states in 1963, and Connecticut in 1965 after the U.S. Supreme Court's decision. *Griswold v. Connecticut* expedited the repeal of anti-obscenity statutes in all the remaining states between 1965 and 1971. In the empirical part, we follow Bailey (2010) and use the period between 1957 and 1965, the year of the U.S. Supreme Court decision, as the period in which the introduction of *Enovid* interacted with the anti-obscenity laws generated what is arguably exogenous variation in cohort size.

We will start by giving some descriptive and graphical evidence on the effect of the source of variation described above on our variables of interest. To do this, we divide the states in two groups: states that enacted sales bans of anti-conception methods and the remaining states. Before presenting the evidence, it is important to emphasize a difference between this paper and Bailey (2010). Bailey is interested in the relationship between the anti-obscenity laws and birth rates of married women after the pill was introduced. She finds that states with the sales ban experienced a marital birth rate that was 8% higher than the remaining states during the period 1957-1965. In this paper we are interested in the link between the anti-obscenity laws and the following three variables: cohort size at birth, cohort size at marriage age, and marriage rates. Differences in the growth in cohort size at birth between the two groups of states provide the first evidence of the effect of the Comstock laws on cohort size between 1957 and 1965. We find that in states with the sales ban, the growth in cohort size at birth was about three percent higher than in states with no ban.

We will now show graphically the relationship between the anti-obscenity laws and cohort size, starting with cohort size at birth followed by cohort size at marriage age. In Figure 6 we report the difference in growth of cohort size at birth between states with the sales ban and the remaining states from 1950 to 1970. We follow Bailey (2010) and present the graphical results separately for the four Census regions. There are two features that are worth discussing. First, after the introduction of the pill, in all regions, states with the ban experienced larger growth in cohort size at birth. The Figure suggests that in the South the states with the ban were reducing the gap in number of births with the states without ban before the introduction of the pill. But it also suggests that the process was expedited by the introduction of the pill. The second noteworthy feature is that in all regions, when states started to outlaw the sales bans on contraceptives, the growth in cohort size at birth in states with the ban started to converge to the growth in states with no ban. The convergence continues until 1965 when the *Griswold v. Connecticut* decision took place, at which point the two groups of states have similar rates of growth in cohort size. These two features taken together generate the hump shape in the difference in growth of number of births that characterizes all census regions.

In Figure 7 we replace cohort size at birth with cohort size at age 25, which represents a measure of cohort size at marriageable age. The figure displays patterns that are similar to the ones observed for cohort size at birth. Between 1957 and 1965, in all regions the difference in growth in cohort size at 25 between the two groups has the familiar hump shape that was observed in the previous graph. This indicates that cohort size at the time of marriage was affected in the expected way by the interaction of the Comstock laws with the introduction of the pill: it increased the growth in size of cohorts born in states with the ban in the period considered until the anti-obscenity laws started to be repealed. The growth in cohort size in these states then started to converge to the growth experienced by states with no ban.

Figure 7 describes graphically the first stage of an IV regression where the introduction of the pill interacted with the sales bans is used as an instrument. We will now formally use that variation in a standard IV setting. To do that, we construct two dummy variables. The first one, $ban_s$, is equal to one for all cohorts from state $s$ that adopted a sales ban on contraceptives and zero otherwise. The second dummy variable, $ban*pill_{c,s}$, takes a value of one if a cohort $c$ was born between 1957 and 1965 in a state $s$ that enacted a contraceptive ban and zero otherwise. Then in the first stage we regress the $n$-year difference in log cohort size at the time of marriage on the two dummy variables and a set of controls, i.e.

$$\log \frac{y_{c,s}}{y_{c-n,s}} = \alpha + \beta_1 ban_s + \beta_2 ban * pill_{c,s} + \sum_{c,r} \pi_{c,r} + X'\gamma + \varepsilon_{c,s}, \tag{1}$$

where $y_{c,s}$ is the size of cohort $c$ in state $s$, $y_{c-n,s}$ is the same variable for cohort $c-n$, $\pi_{c,r}$ are cohort-region fixed effects, and $X'$ is a set of control variables that includes an indicator equal to 1 if the state had a physician exception, the physician indicator interacted with $pill_{c,s}$, and an indicator equal to 1 if the state enacted an advertising ban on contraception.

The results of the first stage are presented in Tables 4, where we report the effect of the anti-obscenity laws on 1-year, 3-year, 5-year, and 7-year differences in log cohort size. Consistent with the graphical evidence provided in Figure 7, we find that after the introduction of the pill and before the repeal of the Comstock laws, the sales ban on contraceptives had a positive and statistically significant effect on cohort size at marriage age in all cases. Consistently with the graphical evidence, the effect increases when we go from a 1-year difference to a 5-year difference. For the 1-year difference, a sales ban increases cohort size by 0.012%, whereas the 5-year differences raises cohort size by 0.041%. The coefficient on the 7-year difference is similar in size. The F-tests to evaluate the strength of the instruments are between 10.11 and 19.22 in our four specifications. An additional result is that the coefficient on $ban_s$ is always small and statistically insignificant suggesting that the sales ban had no effect on cohort size before the introduction of the pill. This finding is consistent with Bailey's results which indicate that the Comstock laws had no effect on other forms of contraception.

In the second stage we use a specification similar to the one employed with the cross-state variation, i.e.

$$\log \frac{mar_{c,s}}{mar_{c-n,s}} = \beta_0 + \beta_1 \log \frac{size_{c,s}}{size_{c-n,s}} + \sum_{c,r} \pi_{c,r} + X'\gamma + \varepsilon_{c,s},$$

except that now we instrument cohort size growth with $ban_s$ and $ban * pill_{c,s}$. Before presenting the results, it is important to remark that to construct the share ever married we must use the decennial Censuses since the CPS does not have enough state-level observations. The decennial Censuses have two limitations. First, in principle the share ever married can be computed for each cohort born in a particular state if one observes in the Census data a recall variable measuring the age at first marriage. Unfortunately, after 1980 this variable is not available in the Censuses. The second limitation is that in each decennial Census we only observe a particular cohort at a particular age. We therefore cannot directly compute the share ever married for each cohort. Instead, we rely on the the following strategy. In each Census, we first consider all individuals between the ages of 25 and 45. We then compute the share ever married for each cohort born in a particular state. Notice that we cannot use this variable directly in our regressions because it is affected by the age at which we observe a particular cohort in a particular Census. To address this issue, we regress the computed share ever married on age, state, cohort, and cohort-region dummies. We then remove the effect of age by subtracting the estimated coefficient on the age dummy multiplied by the dummy itself. Finally, we use the constructed variable in our regressions.

The second stage results are reported in Table 5 for men and women separately. The coefficient estimates have the expected negative sign, are statistically different from zero, and large in magnitude. They indicate that during the period considered a 1% increase in cohort size at marriageable age generated a reduction in marriage rates between 0.24% and 0.44%. The point estimates in the IV regressions are somewhat larger in size than the corresponding estimates using the longitudinal or cross-sectional variation. Because we have to construct the age-adjusted share ever married using individuals observed between the ages of 25 and 45, the coefficients are not directly comparable to the results obtained using the longitudinal and cross-sectional variation for share ever married by 30 or 40. Note, however, that in general we would expect the IV coefficients to be negative and larger in magnitude given the discussion in section 4.4 on potential endogeneity concerns, since most plausible omitted variables would bias the coefficients positively toward zero. We conclude that the IV findings are consistent with the results obtained using longitudinal and cross-state variation and they suggest that there is a causal relationship between changes in cohort size and changes in marriage rates.

We conclude this section with a discussion of a potential threat to our IV strategy. It is possible that the negative relationship between cohort size and share ever married we find in our IV regressions is generated by some type of selection process governing who becomes a mother in states without the ban after the introduction of the pill. The most serious hypothesis that could confound the interpretation of our results is that, after the introduction of the pill, mothers in states without the ban give birth to fewer children who are positively selected along some dimension. If those children are more likely to marry, as the literature suggests, our IV regressions will estimate a negative relationship between our two main variables. To evaluate this hypothesis, we follow Ananat and Hungerman (2012) and test whether children born in states where the pill was banned are more or less likely to have low birth weight. We employ the same specification used in the first stage of the IV estimation except that we use levels instead of differences with two new dependent variables: the share of children born with extremely low birth weight, which is defined as a birth weight below 1500 grams, and the share of children with low birth weight, which is a birth weight below 2500 grams. The estimation results are reported in Table 6. Using both dependent variables, the estimated coefficient on the interaction between the ban and the introduction of the pill is small and statistically insignificant. We find therefore no evidence that the initial access to the pill had an effect on the fraction of children born with low weight. Our result is different from the one obtained in Ananat and Hungerman (2012), where the authors find that initially access to the pill increased the share of children born with low weight. But the sample used is also different. Here, we consider the sample of married women, whereas Ananat and Hungerman (2012) study the behavior of single women younger than 21. The different results can therefore be rationalized by a more widespread early use of the pill by married women relative to young single women.[8]

Ananat and Hungerman (2012) also find weak evidence that early access to the pill had the effect of increasing the share of children born in poor families. To measure that effect, we would need micro

[8] A second difference between our paper and Ananat and Hungerman (2012)'s paper is that we use the interaction between bans on contraception and the introduction of the pill as our main source of exogenous variation, whereas Ananat and Hungerman (2012) use restrictions on access to the pill for minors.

data. Unfortunately, for our sample period the only micro data containing the required variables is the Census, which is only available for the decennial years and therefore cannot be used here. Observe, however, that this negative selection would be a threat to our IV estimates only if children born in low income families are more likely to marry. In this case, the negative selection would generate the negative relationship between cohort size and marriage rates we observe in the data. But the literature on household formation appears to rule out this alternative.[9]

The results discussed in this section indicate that there is a causal negative relationship between cohort size and marriage rates. The rest of the paper proposes and tests a potential mechanism that can explain this relationship.

# 5 A Dynamic Search Model of the Marriage Market

In this section we develop a dynamic search model of the marriage market. We decided to focus on this type of model for two reasons. First, it is one of the simplest models that can match the negative causal relationship between cohort size and marriage rates observed in the data. Second, the search model allows us to easily capture the dynamic nature of the marriage market which is an essential part of the mechanism that enables us to explain the link between cohort size and marriage rates. This explains why we did not consider 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), which is not as well suited to incorporate the dynamic aspects of the marriage market.

In the rest of the section, we first outline the main features of the model. We then show that the model can match the negative causal 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.

## 5.1 Characterization of the Model

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 $\theta^i_{a,t}$. The two spouses then decide whether to marry with the objective of maximizing their lifetime utility.

[9]For instance, the handbook chapter by Black and Devereux (2010) indicates that there is a positive intergenerational correlation in income and education. Moreover, Stevenson and Wolfers (2007) find no difference in the share ever married by education and therefore income. These two results suggest that children of low income parents are not more likely to marry.

We now introduce the main assumption of the model. We assume that women meet men 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. Two ideas form the basis for this assumption. First, women's fertile lifespan is shorter than men's since women are fertile only in their first part of their adult life, whereas men are fertile for most of their adult life. Second, one important benefit of marriage is that it is an effective arrangement for having and raising children. These two ideas imply that, with age, the value of getting married for a woman declines faster than the value for a man. Our assumption that this value for a woman is zero in the second part of her adult life 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 $\delta$, whereas the within-period utility of being married for the couple as a whole will be denoted by $\eta$. 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 old man chooses to be single in the second period, his lifetime utility takes the following form:

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

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

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

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 $\eta$ 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 an old 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, therefore,

$$w_{1,t}^{m}(\eta) = v_{1,t}^{m} + \gamma \left[ v_{0,1,t} - v_{1,t}^{m} - v_{0,t}^{w} \right] = v_{1,t}^{m} + \gamma_{m} \left[ \frac{1-\beta^{T}}{1-\beta} \eta + \beta^{T} \delta - v_{1,t}^{m} - v_{0,t}^{w} \right], \tag{2}$$

where the parameter $\gamma \in [0, 1]$ allows for possible asymmetries in the way the marriage surplus is divided 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 $\gamma$ 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. We can therefore determine the match quality $\eta$ above which the couple will choose to marry. 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 old man before he enters the marriage market. If in period $t$ this man meets a woman and draws a value $\eta$, Nash-bargaining implies that he receives the following share of the couple's lifetime utility:

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

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

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

It is composed of three parts. The first part describes the value for the old man of meeting a woman with a match quality $\eta$ 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 $\eta$ that is below the reservation value $\underline{\eta}_{1,t}$ 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 $\underline{\eta}_{1,t} = 2\delta$, 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:

$$v_{1,t}^m \frac{1-\beta}{1-\beta^T} = \delta + \gamma \left\{ E\left[\eta \left| \eta \geq 2\delta \right.\right] - 2\delta \right\} \left(1 - F\left(2\delta\right)\right) \theta_{1,t}^m. \tag{3}$$

We are now in the position to consider the decision of a young man. He 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

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

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:

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

We can now substitute for the expected value function of an old man using equation (3) and simplify some of the terms to obtain the following equation for the reservation value of a young man:

$$\underline{\eta}_{0,t} = 2\delta + \beta \frac{1-\beta^T}{1-\beta^{T+1}} \gamma \left\{ E\left[\eta \left| \eta \geq 2\delta \right.\right] - 2\delta \right\} \left(1 - F\left(2\delta\right)\right) \theta_{1,t+1}^m. \tag{4}$$

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

Before concluding this subsection, we want to remark that it would be straightforward to add congestion effects to our search model. Their advantage is that, if properly modeled, they mechanically amplify the effects generated by our standard search model. Their main limitation is that they are impossible to test without data on dating behavior which is not available for our sample period. For this reason, we have decided not to incorporate them in our model.

### 5.2 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 young man meets a woman $\theta_{0,t}^m$ and the corresponding probability for an old 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

$$\theta_{0,t}^m = \theta_{1,t}^m = \frac{N_{0,t}^w}{N_{0,t}^m + N_{1,t}^m}. \tag{5}$$

The probability $\theta_{0,t}^m$ is the correct measure in our model of what is called sex-ratio in the literature on household formation. The number of individuals of age 0 is exogenously given by the cohort size of a generation. However, the number of old men in the marriage market $N_{1,t}^m$ is endogenously determined by the decisions of young 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 young men who did not meet a woman at $t-1$ plus the number of young men who met a woman at $t-1$ but draw a match quality $\eta$ lower than the reservation value, i.e.

$$N_{1,t}^m = N_{0,t-1}^m \left(1 - \theta_{0,t-1}^m\right) + N_{0,t-1}^m \theta_{0,t-1}^m F\left(\underline{\eta}_{0,t-1}\right) = N_{0,t-1}^m \left(1 - \theta_{0,t-1}^m \left(1 - F\left(\underline{\eta}_{0,t-1}\right)\right)\right). \tag{6}$$

We can now replace for $\theta_{0,t-1}^m$ using (5) 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 old men in the marriage market $N_{1,t}^m$ are constant over time. We therefore have that

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

We can now solve for $N_1^m$ and obtain

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

Generally, men and women have identical cohort size, i.e. $N^m_{0,t} = N^w_{0,t} = N_{0,t}$.[10] In this case the solution for $N^m_1$ simplifies to

$$N^m_1 = N_0 F(\underline{\eta}_0)^{\frac{1}{2}}.$$

If we substitute $N^m_1$ back into $\theta^m_j$, we have

$$\theta^m_0 = \theta^m_1 = \frac{N^w_0}{N^m_0 + \sqrt{(N^m_0)^2 - N^m_0 N^w_0 + N^m_0 N^w_0 F(\underline{\eta}_0)}}.$$

If men and women have identical cohort size, $\theta^m_j$ simplifies to

$$\theta^m_0 = \theta^m_1 = \frac{N_0}{N_0 + N_0 F(\underline{\eta}_0)^{\frac{1}{2}}} = \frac{1}{1 + F(\underline{\eta}_0)^{\frac{1}{2}}}.$$

To determine the reservation value of young men in steady state, we can substitute for $\theta^m_1$ in the equation that determines the reservation value (4). We can then derive, for the case in which $N^m_0 \neq N^w_0$, the following equation for the steady state reservation value:

$$\underline{\eta}_{ss} = 2\delta + \beta \frac{1-\beta^T}{1-\beta^{T+1}} \gamma \{E[\eta | \eta \geq 2\delta] - 2\delta\} (1 - F(2\delta)) \frac{N^w_0}{N^m_0 + \sqrt{(N^m_0)^2 - N^m_0 N^w_0 + N^m_0 N^w_0 F(\underline{\eta}_{ss})}},$$

If $N^m_0 = N^w_0$, the equation simplifies as follows:

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

Note that $F(\underline{\eta})$ 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 $N^m_0$ and $N^w_0$. The following Proposition summarizes the result.

**Proposition 1** *In steady state, there is a unique reservation value for marriage* $\underline{\eta}_{ss}$. *It does not depend on cohort size if* $N^m_0 = N^w_0$.

## 5.3 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

[10] 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 incarcerated during their marriage years. As a consequence, the relevant cohort size for African-American men is smaller than the corresponding cohort size for women.

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 incarceration and mortality rates are higher for men of marriageable 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 cohort size from $N_0$ to $N_0 + \Delta$. 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 (5), 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 old 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 $\theta_{j,t}^m$, we have that in period $\tau$

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

The previous equation implies that a positive cohort shock $\Delta$ increases the probability that a man of any age meets a woman, whereas a negative cohort shock has the opposite effect. In our economy there are always more men than women in the marriage market. As a consequence, the probability that a woman meets a young man, $\theta_t^w = \frac{N_{0,t}}{N_{0,t}+N_{1,t}^m}$, 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 young men.

We can now determine the effect of a shock to cohort size on the reservation value of young men $\underline{\eta}_{0,\tau}$. Notice that in the determination of $\underline{\eta}_{0,\tau}$ a young man compares the value of getting married at $\tau$ 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 old men at $\tau + 1$, which can be written as follows:

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

Using equation (6), we can substitute for $N_{1,\tau+1}$ to obtain the following expression:

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

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

$$\underline{\eta}_{0,\tau} = 2\delta + \beta \frac{1 - \beta^{T}}{1 - \beta^{T+1}} \gamma \left\{E\left[\eta \left| \eta \geq 2\delta\right.\right] - 2\delta\right\}\left(1 - F\left(2\delta\right)\right) \frac{1}{1 + \left(1 - \theta_{0,\tau}^{m}\left(1 - F\left(\underline{\eta}_{0,\tau}\right)\right)\right)}. \tag{7}$$

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

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

Earlier in this section we have shown that, with a positive shock to cohort size, $\theta_{0,\tau}^{m} > \theta_{0,ss}^{m}$. 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 and permanent shock to cohort size in period $\tau$ 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 and permanent shock to cohort size in period $\tau$ reduces the fraction of cohort $\tau$'s women who get married by increasing the reservation value of marriage of young couples. A negative shock in period $\tau$ has the opposite effect.*

**Proof.** In the appendix. ■

To provide the insight behind this result, consider an increase in cohort size. After this event, old men become a scarce resource. This change has two effects. First, the fraction of women who marry mechanically declines because women are now less likely to meet old men, who have lower reservation values and, hence, higher probability of marriage. Second, young 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 old men because there are relatively fewer of them. Since the reservation value of old 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 young men because of the increase in cohort size. For a constant reservation value for young men, this effect would decrease the average age difference at marriage. However, young men become more selective in their decision to marry. As a result, fewer of the young men that meet a woman end up getting married. The combined effect of young men on the average age difference is therefore unclear. Proposition 4 shows that the effect generated by old men and the effect produced by a greater probability of meeting a young men dominate the effect of a higher reservation value. The increase in cohort size therefore reduces the average age difference at marriage.

Proposition 4 enables us to directly test the mechanism used by our model to generate the negative relationship between cohort size and marriage rates observed in the data. In our model the reduction in age difference between spouses subsequent to an increase in cohort size is generated by the change in the probability that a man meets a woman, $\theta_{0,t}^{m}$, or equivalently by a change in the sex ratio, which is at the core of our search model. It is difficult to think of an alternative well-specified and testable model that generates the same implication. The testable implication contained in Proposition 4 is therefore informative about the ability of our model to characterize the marriage market.

# 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 white individuals because, as argued in the previous section, the assumption $N_0^m = N_0^w = N_0$ is not satisfied in the data for the black population.

## 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 construct the variable age difference between spouses as follows. First, we consider all married women of a given age. We then compute the difference between their age and the age of their spouse. Finally, we calculate the average for each cohort. The Data Appendix B provide details on the data sets and ages of married women used to construct this variable. It is worth remarking that we cannot perform the test using the introduction of the pill as an instrument. The reason for this is that the age difference cannot be constructed consistently for all cohorts born between 1945 to 1970 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.

In Figure 8 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 8 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. In our model, the change in age difference at marriage is generated by changes in the sex ratio interacted with the decision about whether to marry. A decline in age difference should therefore be accompanied by an increase in the sex ratio. To determine whether that is the case, in Figure 8 we also report the evolution of the sex ratio. This variable is constructed by dividing the number of young women by the number of young men plus the number of old men, where young women are all women between the ages of 18 and 29, young men are all men in the same age range, and old men are all men between the ages of 30 and 40. Given the way our variable is constructed, the first available observation is for 1932. The sex ratio evolves over time in a way that is consistent with our model. An increase in cohort size is accompanied by a rise in sex ratio and by a decline in the age difference between spouses.

The first column of Table 7 reports the coefficient estimate 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. To avoid mixing observations from differenct data sets, we only consider cohorts born in 1930 or after which are all observed in the CPS. 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.

In the second column of the same table we report the coefficient estimate obtained using cross-state variation. We follow a similar approach to the one used in the regressions estimating the effect of cohort size on the share ever married and instrument the changes in cohort size at adulthood with changes in cohort size at birth. The estimated coefficients for the decennial cohorts born between 1910 and 1970 are consistent with the search model. Our estimated coefficient is negative and statistically significant indicating that an increase in cohort size has a negative effect on the age difference between spouses. 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 enables us to additionally test 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 $\eta$. We assume that it is distributed according to a beta distribution with shape parameters $\alpha_1$ and $\alpha_2$ 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 $\delta$ varies over time according to the following equation:

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

where $\nu_t$ is drawn from a uniform distribution defined on the interval $[-0.2, 0.2]$.[11]

The third set of assumptions we make are related to the lifespan of an individual. In practice, individuals from a given cohort will participate in the marriage market 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 multiple 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, five periods. To implement the assumption that each 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, consecutive increases in cohort size will have a larger effect than a single increase, because the newcomers compete in a marriage market that is progressively more crowded. A similar argument applies to declines in cohort size.[12]

As additional assumptions, we set the annual discount factor equal to 0.98 and consider a symmetric Nash-bargaining by setting $\gamma$ 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 young or old man meets a woman. Specifically, these probabilities now take the following form:

$$\theta^m_{0,t} = \theta^m_{1,t} = \frac{N^w_{0,t}}{\left(N^m_{0,t} + N^m_{1,t}\right)\phi}.$$

The probability that a woman meets a young or old man does not change since the parameter $\phi$ 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 $\delta$, which is assumed to be identical across gender and over time; the two shape parameters of the beta distribution $\alpha_1$ and $\alpha_2$; the fraction of individuals that are unwilling to marry $1 - \phi$. These parameters are 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.

In the estimation, we use as our set of moments the fraction of women never married in a cohort

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

[12] 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.

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 $\delta$ and the parameter that determines the fraction of men who are unwilling to marry $\phi$. To see this, observe that that the value of being single $\delta$ is linked to the fraction of individuals never married in a given cohort. Everything else equal, a higher value of $\delta$ increases the share of individuals who choose to stay single in each cohort. The parameter $\phi$ enters equation (4) 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 $\phi$: a larger value for $\phi$ generates a smaller slope. As a consequence, changes in the meeting probabilities will have smaller effects on the reservation value of young men and therefore on marriage rates if $\phi$ is larger. In the model, changes in meeting probabilities are mainly generated by variations in cohort size. The parameter $\phi$ 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 $\delta$, $\alpha_1$, $\alpha_2$, and $\phi$ 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.[13]

The estimated parameters are reported in Table 8. Our estimates of $\alpha_1$ and $\alpha_2$ are 0.020 and 0.072. These values imply that the mean of the distribution is equal to 0.217, the mode is equal 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 old man marry only if $\eta > 2\delta$. Given our estimate of $\delta$, 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 old men marry only if they draw a relatively high value for match quality. In addition, observe that young men have a higher reservation value. Consequently, women paired with young men will also marry only if their match quality is relatively high. Finally, our model rationalizes the data by estimating that 13.3% of the

[13] We have proven that the parameter $\alpha_1$ can be identified using as a moment the probability that an old man marries plus the corresponding probability for a woman; $\alpha_2$ can be identified using as a moment the probability that a young man marries divided by the probability that a woman marries times the previous moment; the parameter $\phi$ 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 old man does not marry; the parameter $\gamma$ 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.

men in the marriage market in a given period 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 9 the marriage rates observed in the data jointly with the marriage rates simulated using the estimated model, both as a function of cohort. Figure 9 shows that the model can quantitatively replicate the medium and long run variation in the share of never married across cohorts observed in the data. It is noteworthy that the search model can generate the big change in the share never married observed in the data between the early thirties and the late seventies, an increase from about about 5 percent in the thirties to about 17 percent seventies. This outcome is particularly remarkable since our model is very parsimonious. We only have four parameters to match 51 moments.

We conclude this section with an important remark. In this paper, we have provided evidence that positive changes in cohort size cause a decline in marriage rates and that negative changes have the opposite effect. We have also proposed a possible mechanism which cannot be rejected using our tests. However, as it is standard when a particular model is tested, we are not able to directly reject that changes in cohort size cause marriage rates to vary through different mechanisms. One possibility is that an increase in cohort size reduces income of that cohort and the reduction in income generates the decline in marriage rates. To evaluate this alternative, we can only rely on the evidence provided in other papers and described in section 2, which suggests that changes in income cannot explain changes in marriage rates for some of the periods we investigate in this paper. A thorough investigation of the effect of changes in cohort size on other economic and social variables is important, but it is left for future research.

## 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 cross-state variation in the adoption of the pill we provide evidence in support of the following two results. First, cohort size can explain on its own more than 50% of the variation in U.S. marriage rates. Second, an increase in cohort size reduces marriage rates and a decrease has the opposite effect. We cannot rule out the hypothesis that cohort size affects marriage rates through a third variable - for instance, income. We leave this for future research. However, if one believes that the cross-state variation in the introduction of the pill is exogenous, in the paper we have provided evidence that there is a causal relationship between cohort size and marriage rates.

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 a positive change in cohort size reduces the age difference between spouses, and a negative change increases it. Finally, we test the model in two different ways. We first test whether the derived implication can be rejected. In the data, an increase 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 that we document. The estimated model can match the long-run variation in marriage rates observed in the data.

Our results have important implications for policy analysis. In recent years, politicians and policy makers have begun to consider and implement policies that increase the fraction of married individuals with the intent of reducing the poverty rate. Several examples of such policies exist. For instance, Temporary Assistance to Needy Families (TANF) allows states to use a fraction of its funds to implement policies designed at increasing the share of married individuals. In West Virginia, households receiving TANF receive an additional 100 dollars a month if they are headed by a legally married household. States have also undertaken other policies to promote marriage. Former Utah governor Leavitt declared a "Marriage Awareness" week and created a commission whose goal was to develop strategies to promote marriage. "First Things First" in Tennessee and "Healthy Marriages Grand Rapids" in Michigan are other examples of programs designed to promote marriage. Our findings indicate that these types of policies may at best only have short term effects, since in the medium and long run marriage rates are for the most part outside the control of policy makers and politicians.

## 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.

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.

Ananat, Elizabeth Oltmans, and Daniel M. Hungerman. 2012. "The Power of the Pill for the Next Generation: Oral Contraception's Effects on Fertility, Abortion, and Maternal and Child Characteristics." *The Review of Economics and Statistics* 94 (1): pp. 37–51.

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.

Angrist, Joshua D., and Victor Lavy. 1996. "The Effect of Teen Childbearing and Single Parenthood on Childhood Disabilities and Progress in School." *NBER Working Paper*, no. 5807.

Bailey, Martha J. 2010. ""Momma's Got the Pill": How Anthony Comstock and Griswold v. Connecticut Shaped US Childbearing." *The American Economic Review* 100 (1): pp. 98–129.

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*.

Black, Sandra E., and Paul J. Devereux. 2010. "Recent Developments in Intergenerational Mobility." *Prepared for the Handbook in Labor Economics 2010*.

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-Andre, 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.

Gruber, Jonathan. 2004. “Is Making Divorce Easier Bad for Children? The Long-Run Implications of Unilateral Divorce.” *Journal of Labor Economics* 22 (4): pp. 799–833.

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.

Killingsworth, Mark R., and James J. Heckman. 1986. “Female Labor Supply: a Survey.” *Handbook of Labor Economics, Edited by O. Ashenfelter and R. Layard,* 1:pp. 103–204.

King, Miriam, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick. 2010. *Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database].* Minneapolis: University of Minnesota.

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.

McLanahan, Sara, and Christine Percheski. 2008. "Family Structure and the Reproduction of Inequalities." *The Annual Review of Sociology* 34:pp. 257–76.

Moffitt, Robert A. 2000. "Welfare Benefits and Female Headship in U.S. Time Series." *The American Economic Review* 90 (2): pp. 373–377.

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

Neal, Derek. 2004. "The Relationship between Marriage Market Prospects and Never-Married Motherhood." *The Journal of Human Resources* 39 (4): pp. 938–957.

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.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. *Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database].* Minneapolis: University of Minnesota.

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

Seefeld, Kristin S., and Pamela J. Smock. 2004. "Marriage on the Public Policy Agenda: What Do Policy Makers Need to Know from Research?" *PSC Research Report, University of Michigan*, no. 04-554.

Stevenson, Betsey, and Justin Wolfers. 2007. "Marriage and Divorce: Changes and Their Driving Forces." *The Journal of Economic Perspectives* 21 (2): pp. 27–52.

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

Wolfers, Justin. 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 (8), 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 (7) 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 young man times the probability she marries him plus the probability she meets an old 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

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

### 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 young man times the corresponding probability plus the difference in age conditional on marrying an old man times the corresponding probability, i.e.

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

Without loss of generality suppose the difference in age if a woman marries a young 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, $E\left[\Delta age\right]$ can be written as follows:

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

Notice that $y_1$, $z$, and $1-F\left(2\delta\right)$ 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 old 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 $\Delta$ and suppose it reduces $N_0(1-F(x))$. We will show that, if this is true, the number of old men in the period that follows the shock to cohort size is so large that the reservation value of young 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 $\theta'$ and $x'$ the probability that a man meets a woman and the reservation value of a young 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 young men who in this period do not meet a woman plus the number of young men who meet a woman and decide not to marry. Hence,

$$\begin{aligned} N_1' - N_1 &= (N_0+\Delta)\left(1-\theta'\right) + (N_0+\Delta)\theta' F\left(x'\right) - N_0(1-\theta) - N_0\theta F(x) \\ &= -(N_0+\Delta)\theta'\left(1-F\left(x'\right)\right) + N_0\theta(1-F(x)) + N_0 + \Delta - N_0 \\ &= -(N_0+\Delta)\theta'\left(1-F\left(x'\right)\right) + N_0\theta'(1-F(x)) + N_0\left(\theta-\theta'\right)(1-F(x)) + \Delta \\ &= -\theta'\left((N_0+\Delta)\left(1-F\left(x'\right)\right) - N_0(1-F(x))\right) + N_0\left(\theta-\theta'\right)(1-F(x)) + \Delta \\ &\geq N_0\left(\theta-\theta'\right)(1-F(x)) + \Delta, \end{aligned} \tag{9}$$

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 $\theta''$, the probability that a man meets a woman the period after the shock takes place, is lower than $\theta$. 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 (9) 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 $\frac{N_1 + N_0(\theta - \theta')(1 - F(x)) + \Delta}{N_0 + \Delta} > \frac{N_1}{N_0}$ if

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

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

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

Hence, we have the desired inequality if

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

or equivalently,

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

Some of the terms cancel out producing the following inequality:

$$N_0 + \Delta > N_0 F(x) + \Delta F(x),$$

which is equivalent to

$$1 > F(x),$$

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:

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

The first term represents the value of meeting a woman with a match quality $\eta$ higher than the reservation value times the probability of this event. The second term describes the value of meeting a women characterized by an $\eta$ 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 young and old men. As a consequence, it takes the following more complex form:

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

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

## B Data Appendix

Table 9 provides a summary of the datasets employed in the construction of the main variables of interest. In the rest of the appendix, we give additional details about how we construct the variables cohort size at birth, cohort size at marriageable age, share ever married, and age differences of spouses.

In the paper we use two different measures of cohort size: cohort size at birth and cohort size at marriageable age. Cohort size at birth is used in three ways: as the main independent variable when we employ longitudinal variation; as an instrument for cohort size at marriage age in the cross-state regressions; and as one of the variables used to determine the effect of the introduction of the pill in states with different anti-obscenity laws. With longitudinal variation we use cohort size at birth as the main independent variable instead of cohort size at marriageable age for two reasons. First, as shown in Figure 1, when cohort size is computed for the U.S. population there is little difference between cohort size at birth and cohort size at in adulthood, since migration from and to the U.S. was limited. Second, we can construct the variable cohort size at birth for cohorts born in 1909 and after. The variable cohort size at marriageable age can only be constructed for cohorts born after the 1940s. By using cohort size at birth we can therefore consider a larger number of cohorts without significant effect on the analysis. As indicated in Table 9, in the longitudinal analysis cohort size at birth is constructed using the U.S. Vital Statistics which provide information on this variable by race from 1909 to 1980. For the cross-state regressions, there are two data sets that can be used to measure cohort size at birth: the U.S. Vital Statistics which record births by race and by state from 1940; and the decennial Censuses which provide information on population counts from the beginning of the twentieth century to 2010. In the cross-state regressions, we work with the decennial cohorts 1910-1970. For consistency, rather than combining two different datasets, we use the Censuses over the entire period of interest. One limitation of the decennial Censuses is that population counts are published for 5-year age groups. From each decennial Census, we therefore record the number individuals between the ages 0 to 4 and we use it to construct the cohort size at birth. Our results do not change if we use data from the U.S. Vital Statistics for the 1940 to 1970 cohorts. In the regressions that use the introduction of the pill as an instrumental variable, we consider cohorts born between 1945 and 1970. We can therefore use the U.S. Vital Statistics to compute cohort size at birth for all of them.

Cohort size at marriageable age is used as the main independent variable in the cross-state regressions and in the regressions that use the introduction of the pill as an instrument. There are two datasets that can be used to measure this variable: the decennial Censuses and the SEER population estimates. SEER records cohort sizes at different ages starting from 1969. Hence, using this dataset we can construct cohort size at marriage age only for some of the decennial cohorts born between 1910 and 1970, which are the ones we consider in the cross-state regressions. For consistency, we therefore construct cohort size at marriage age by recording the number of individuals between the ages 20 and

24 in the decennial Censuses 1930-1990. We also experimented with the 5-year age group 30-34 with similar results. In the regressions that use the pill and the anti-obscenity laws as instruments, we use cohorts born between 1945 and 1970 which are all observed in SEER at age 25 or older. We therefore measure cohort size at marriage using the information in SEER at age 25.

The variable share ever marriage is constructed using a different procedure depending on whether we use longitudinal or cross-state variation. With longitudinal variation, we employ a combination of the CPS, which covers the period 1962-2011, and of the decennial Censuses. In the CPS, we observe the age and the marital status of each respondent. We can therefore easily compute the share ever married by age 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 30 and 35 for different cohorts by considering all individuals who in a given Census are between the ages of 30 and 45. We use a maximum cutoff age of 45 to avoid potential 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. With cross-state variation, for all cohorts we only use information from the Censuses, as sample sizes in the CPS are too small to provide reliable estimates at the state level. In the longitudinal as well as in the cross-state variation, we cannot construct the share ever married for cohorts born before 1914 because the 1960 Census is the first one that records the age at first marriage.

We use the variable "Relationship to household head" in the Census and CPS to record households in which a cohabiting partner is present. The Census began recording unmarried partners only in 1990, and the CPS only in 1995. As a result, in the longitudinal analysis we may miss cohabitations for cohorts born before 1965 when we use 30 as the age cutoff, or 1955 when we use 40 as the cutoff. In the cross-sectional analysis, we may similarly miss cohabitations for cohorts born before 1960 or 1950, depending on the age cutoff. In the data we observe that cohabitation for early cohorts is limited. For the 1965 cohort, the share of individuals cohabiting at age 30 was 2.8% For the 1955 cohort, the share cohabiting at age 40 was 0.76%. We examined data in the National Survey of Families and Households (NSFH) to test whether we miss a substantial number of cohabitations for the cohorts for which we do not have the cohabitation variable, especially at the lower age cutoffs. The first wave of the NSFH (1987-1988) is nationally representative and provides retrospective data on marriage and cohabitation. We use the dataset to examine cohabitation patterns at age 30 for cohorts born 1957 or earlier. We found that cohabitation at age 30 is almost non-existent for pre-baby boom cohorts. From 1945 to 1957, the average share of individuals cohabiting is 0.5%. We conclude that we only marginally underestimate the share ever married or cohabiting at age 30 for the early baby boom cohorts.

The variable age difference between spouses is used to test the search model using both longitudinal and cross-state variation. With longitudinal variation, we use cohorts born between 1930 and 1975 so that we can use the CPS to construct this variable. 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 cross-state variation, the average age difference is computed using the 1940-2000 Censuses, since the CPS does not have enough observations at the state level. In this case, for each decennial cohort born between 1910 and 2000 we consider all women of age 30 to 35 who are married, compute the age difference with their spouse, and calculate the average at the state level.

# C 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 | -0.294*** | -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%. Notes: Newey-West standard errors in parentheses. Each coefficient is the outcome of a separate regression. Regressions include cohorts born after 1914 until the most recent cohort observed at a given age in 2011. The number of observations in each regression is equal to 68 for the share ever married by 30, 63 for the share ever married by 35, and 58 for the share ever married by 40. Sources: IPUMS CPS 1962-2011, IPUMS Census 1960-1970.

**Table 2:** Regression of Change in Log Share Ever Married on Change in Log Cumulative Cohort Size

| | 1-Yr. | 2-Yr. | 3-Yr | 4-Yr | 5-Yr. | 7-Yr. | 10-Yr. |
|---|---|---|---|---|---|---|---|
| Men | 0.049 | 0.017 | -0.157** | -0.220*** | -0.233*** | -0.238*** | -0.277*** |
| By Age 30 | (0.037) | (0.062) | (0.061) | (0.049) | (0.045) | (0.040) | (0.047) |
| | | | | | | | |
| Men | -0.054 | -0.057 | -0.090** | -0.101*** | -0.098** | -0.115*** | -0.163*** |
| By Age 35 | (0.054) | (0.044) | (0.040) | (0.034) | (0.026) | (0.023) | (0.017) |
| | | | | | | | |
| Men | 0.073*** | 0.037 | -0.018 | -0.048** | -0.054** | -0.052** | -0.085*** |
| By Age 40 | (0.016) | (0.038) | (0.029) | (0.022) | (0.023) | (0.023) | (0.021) |
| | | | | | | | |
| Women | -0.064*** | -0.077*** | -0.101*** | -0.095*** | -0.103*** | -0.125*** | -0.156*** |
| By Age 30 | (0.024) | (0.037) | (0.014) | (0.019) | (0.021) | (0.024) | (0.028) |
| | | | | | | | |
| Women | -0.005 | -0.053** | -0.070*** | -0.084*** | -0.081*** | -0.090*** | -0.113*** |
| By Age 35 | (0.056) | (0.021) | (0.011) | (0.012) | (0.015) | (0.016) | (0.015) |
| | | | | | | | |
| Women | -0.011 | -0.005 | -0.008 | -0.022* | -0.036** | -0.056*** | -0.073*** |
| By Age 40 | (0.021) | (0.021) | (0.012) | (0.013) | (0.014) | (0.010) | (0.006) |

* Significant at 10%. ** Significant at 5%. *** Significant at 1%. See notes in Table 1. Newey-West standard errors in parentheses.

**Table 3:** Cross-Sectional Regression of Log Share Ever Married by 30 or 40

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

| 10-Yr. Dfference in Log Cohort Size | OLS | IV (1) | IV (2) |
|---|---|---|---|
| Men | -0.041** | -0.123*** | -0.168*** |
| By Age 30 | (0.019) | (0.041) | (0.056) |
| $R^2$ | 0.807 | 0.791 | 0.819 |
| | | | |
| Men | -0.029*** | -0.047** | -0.040 |
| By Age 40 | (0.010) | (0.018) | (0.028) |
| $R^2$ | 0.555 | 0.551 | 0.595 |
| | | | |
| Women | -0.041*** | -0.104*** | -0.115*** |
| By Age 30 | (0.010) | (0.023) | (0.036) |
| $R^2$ | 0.555 | 0.765 | 0.796 |
| | | | |
| Women | -0.032*** | -0.058*** | -0.048** |
| By Age 40 | (0.010) | (0.017) | (0.023) |
| $R^2$ | 0.539 | 0.528 | 0.605 |
| First Stage Results | | | |
| Log Cohort Size at Birth | | | 0.440*** |
| | | | (0.073) |
| F-test | | | 36.44 |
| $R^2$ | | | 0.831 |

* Significant at 1%. ** Significant at 5%. *** Significant at 1%. Notes: Robust standard errors in parentheses. $N = 288$. Each coefficient is the outcome of a separate, population-weighted regression. We control for cohort fixed effects, and for cohort-region fixed effects in IV-(2). Regressions are for decennial cohorts born between 1910 and 1970. Data includes all states except Hawaii and Alaska. Sources: US Census Population Counts, 1910-1990; IPUMS USA, 1940-2010.

**Table 4:** Comstock Laws and N-Year Differences in Log Cohort Size

Dependent Variable: N-Yr. Difference in Log Cohort Size

| | 1-yr | 3-yr | 5-yr | 7-yr |
|---|---|---|---|---|
| $Ban * Pill_{c,s}$ | 0.012** | 0.035*** | 0.041*** | 0.040** |
| | (0.005) | (0.011) | (0.015) | (0.017) |
| $Ban_s$ | -0.002 | -0.003 | 0.006 | 0.018 |
| | (0.005) | (0.010) | (0.012) | (0.014) |
| N | 1248 | 1152 | 1056 | 960 |

* Significant at 10%. ** Significant at 5%. *** Significant at 1%. Notes: Robust standard errors in parentheses. Regressions are weighted by population and include controls for physician exception, physician exception interacted with "pill", advertising bans, and cohort-region fixed effects. Source: IPUMS USA, 1980-2000. NIH SEER Population Counts.

**Table 5:** N-Year Differences in Log Share Ever Married and N-Year Differences in Log Cohort Size

Dependent Variable: N-Yr. Difference in Log Share Ever Married (Men)

| | 1-yr | 3-yr | 5-yr | 7-yr |
|---|---|---|---|---|
| N-yr Difference | -0.389* | -0.241** | -0.239*** | -0.260*** |
| in Log Cohort Size | (0.206) | (0.101) | (0.091) | (0.080) |
| N | 1248 | 1152 | 1056 | 960 |

Dependent Variable: N-Yr. Difference in Log Share Ever Married (Women)

| | 1-yr | 3-yr | 5-yr | 7-yr |
|---|---|---|---|---|
| N-yr Difference | -0.441*** | -0.311*** | -0.302*** | -0.304*** |
| in Log Cohort Size | (0.161) | (0.093) | (0.0845 | (0.076) |
| N | 1248 | 1152 | 1056 | 960 |

* See note in Table 4.

**Table 6:** Comstock Laws and Birth Weight

Dependent Variable: Birth Weight

| | Share with Birth Weight < 1500 | Share with Birth Weight < 2500 |
|---|---|---|
| $Ban * Pill_{c,s}$ | 0.0001025 | -0.0004241 |
| | (.0002037) | (.0007311) |
| $Ban_s$ | -0.000462 | 0.0006242 |
| | (0.0003064) | (0.0006155) |
| N | 1006 | 1006 |

* See note in Table 4.

**Table 7:** Regressions: Log Age Difference and Log Cohort Size

| | Time Series, 1930-1975 | Cross-State, 1910-1970 |
|---|---|---|
| Log Cohort Size | -0.592*** | |
| | (0.043) | |
| 10-Yr. Difference in Log Cohort Size | | -0.241*** |
| | | (0.092) |
| N | 46 | 288 |
| R-squared | 0.81 | |

*** Significant at 1%. Notes: Robust standard errors in parentheses. The cross-state IV regression controls for cohort-region fixed effects; data includes all states except Hawaii and Alaska. Sources: US Census Population Counts, 1910-1970; IPUMS USA, 1940-2000.

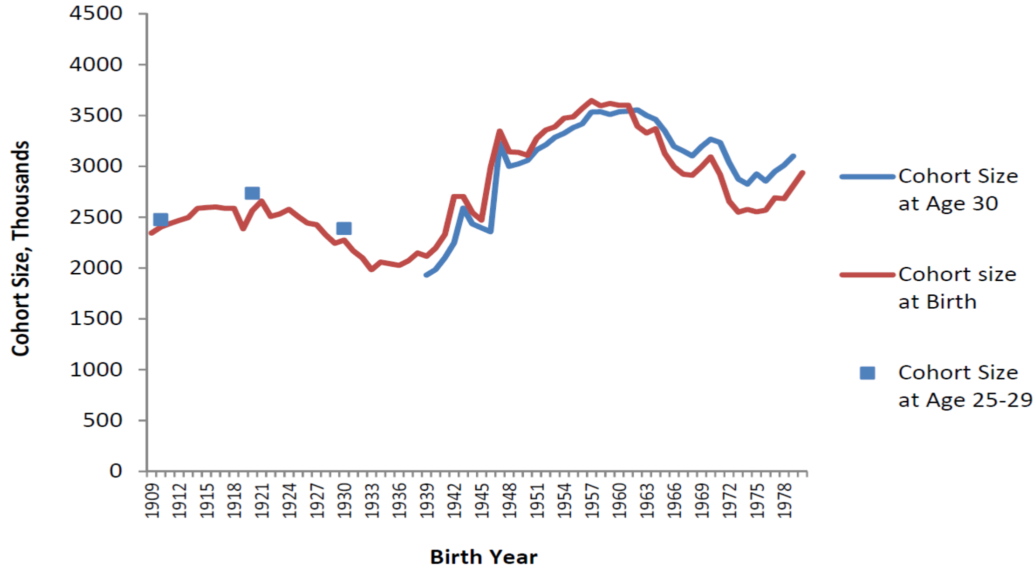
**Table 8:** Estimated Parameters

| 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] |

**Table 9:** Data Sets Used in the Construction of the Main Variables

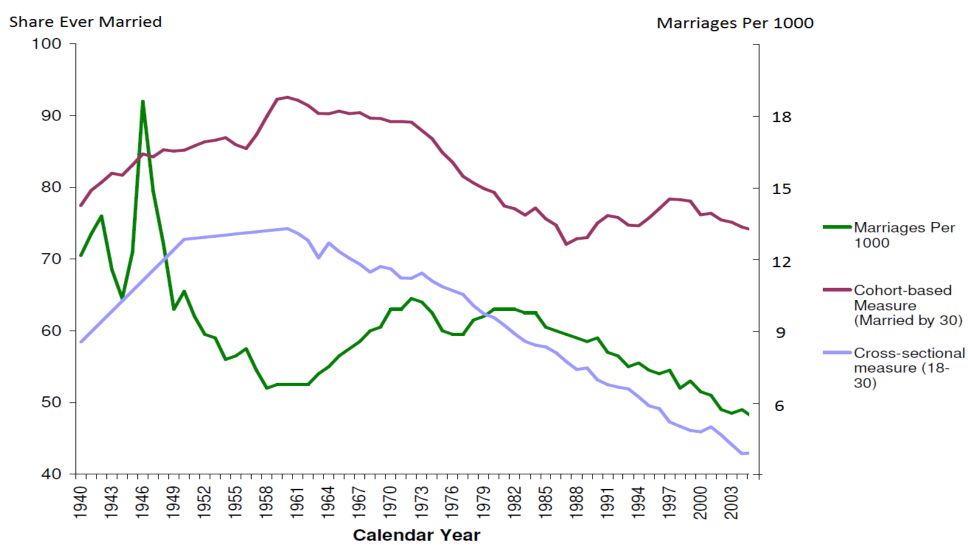
| Variable | Variation of Interest | Dataset |
|---|---|---|
| Cohort Size at Birth | National, Yearly | Vital Statistics of the United States (1909-1980) |
| | State Level, Decennial Years | U.S. Decennial Census, 1910-1970 |
| Cohort Size at Adulthood | National and State Level, Yearly (Around the Introduction of the Pill) | NIH/SEER Population Estimates (1969-2010) |
| | State-Level, Decennial Years | U.S. Decennial Census, 1930-1990 |
| Share Ever Married | National, Yearly | CPS (1962-2011), U.S. Decennial Census, 1960-1970 |
| | State Level, Decennial Years | U.S. Decennial Census, 1940-2000 |
| Age Differences of Spouses | National, Yearly | CPS (1962-2011), U.S. Decennial Census, 1960 |
| | State Level, Decennial Years | U.S. Decennial Census, 1970-2000 |

**Figure 1:** Cohort Size at Birth and at Age 30



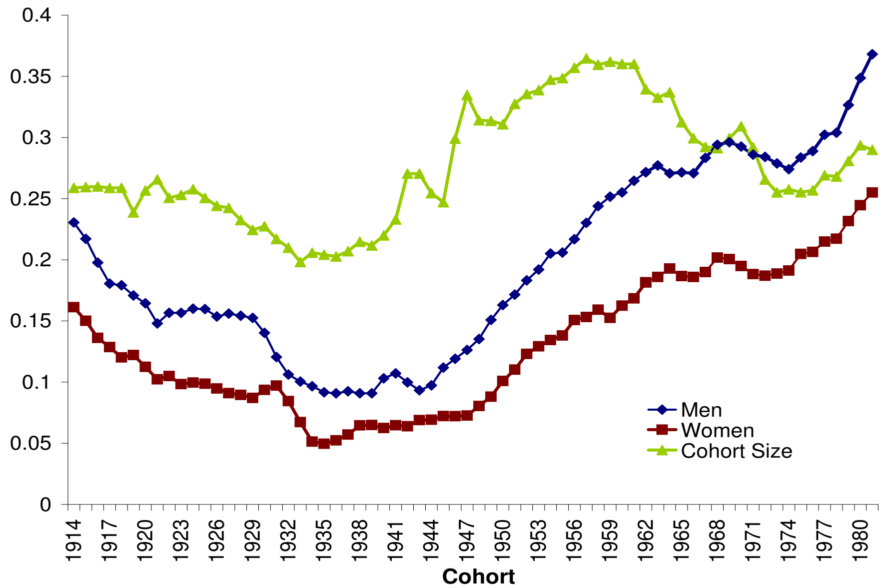

White individuals only. Sources: Vital Statistics of the U.S.; NIH SEER population estimates (1969-2010); U.S. Census 1940-1960.

**Figure 2:** Different Measures of Marriage Rates

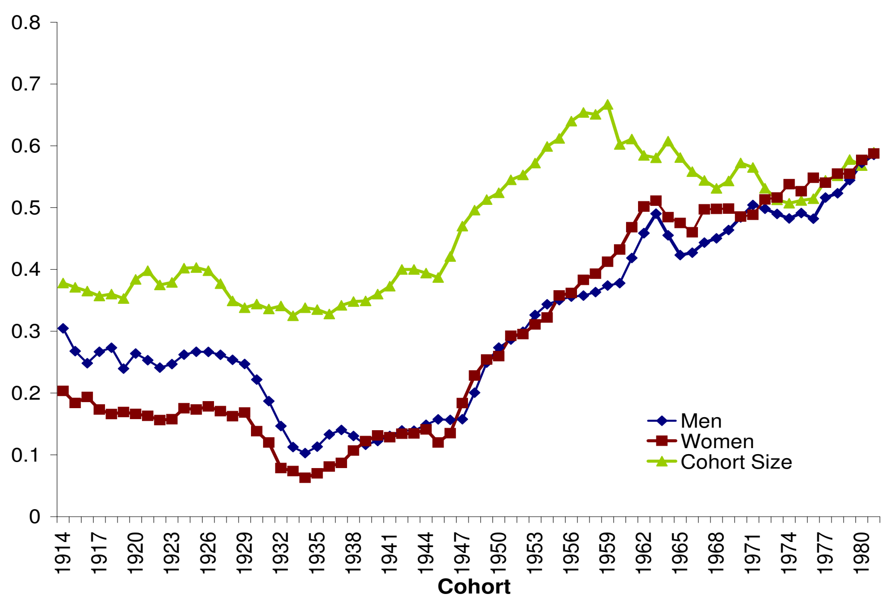



Sources: Vital Statistics of the United States; IPUMS CPS, 1962-2005; IPUMS USA, 1940-1960. The left vertical axis marks the percentage of individuals ever married for the cohort-based and cross-sectional measures; the right axis corresponds to marriages per thousand. To facilitate direct comparison,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

**Figure 3:** Share Never Married By 30



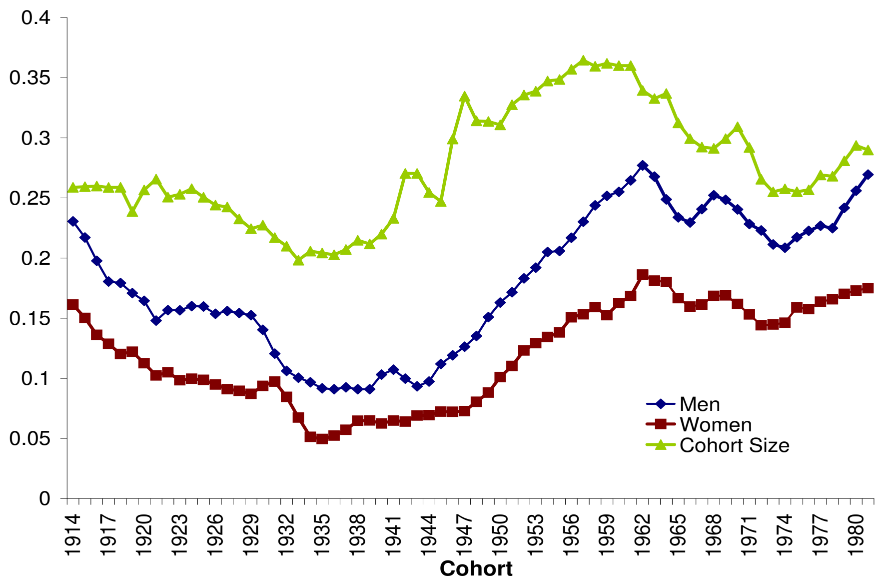

(a) White



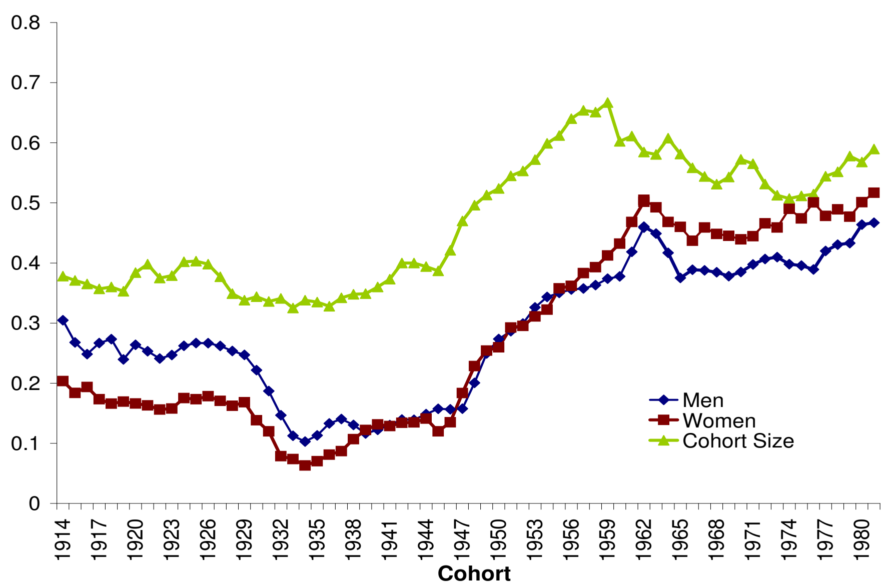

(b) Black

Notes: The vertical axis represents both the percentage of individuals ever married as well as normalized cohort size. In Panel A we normalize cohort size by dividing by 10,000,000; in Panel B by 1,000,000. For share ever married, we graph three-year moving averages. Sources: Vital Statistics of the United States; IPUMS CPS, 1962-2011; IPUMS USA, 1960-1970.

**Figure 4:** Share Never Married and Not Cohabiting By 30



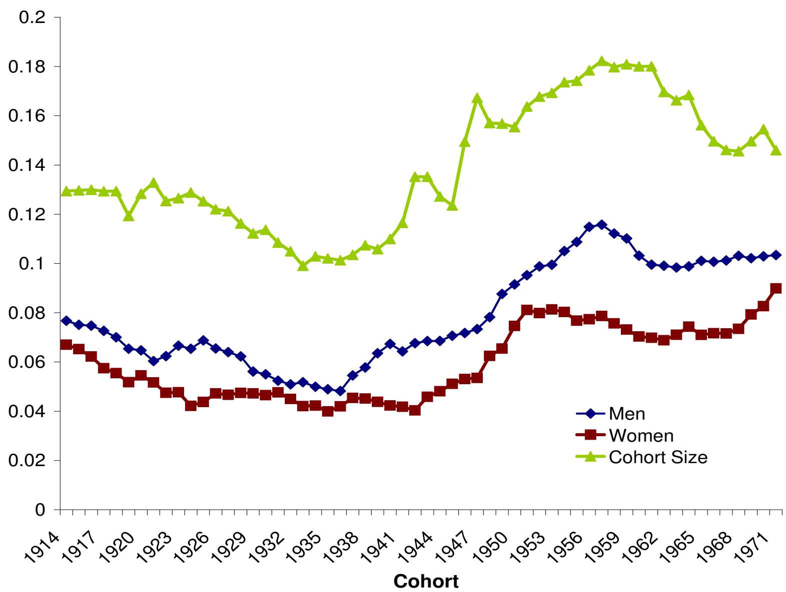

(a) White



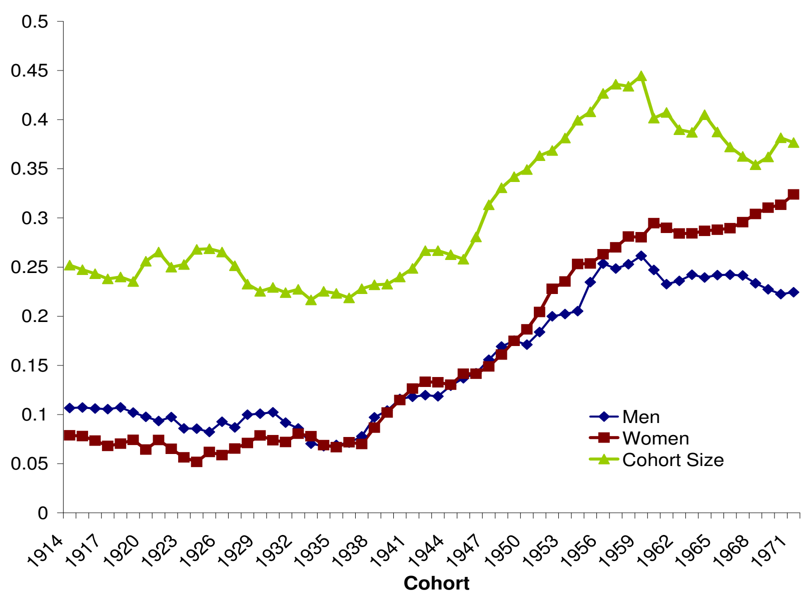

(b) Black

* See note in Figure 3.

**Figure 5:** Share Never Married and Not Cohabiting By 40



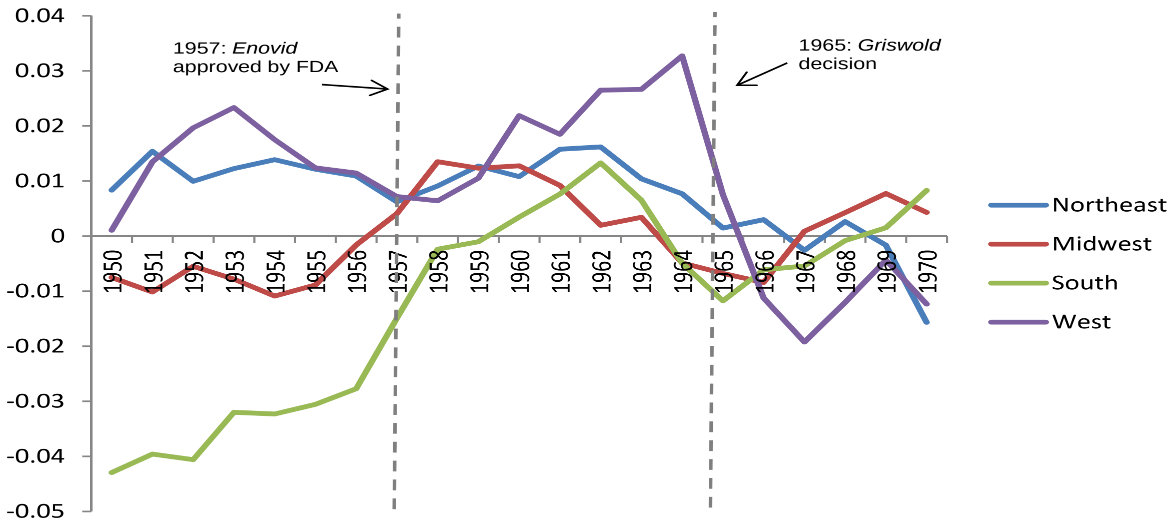

(a) White




(b) Black

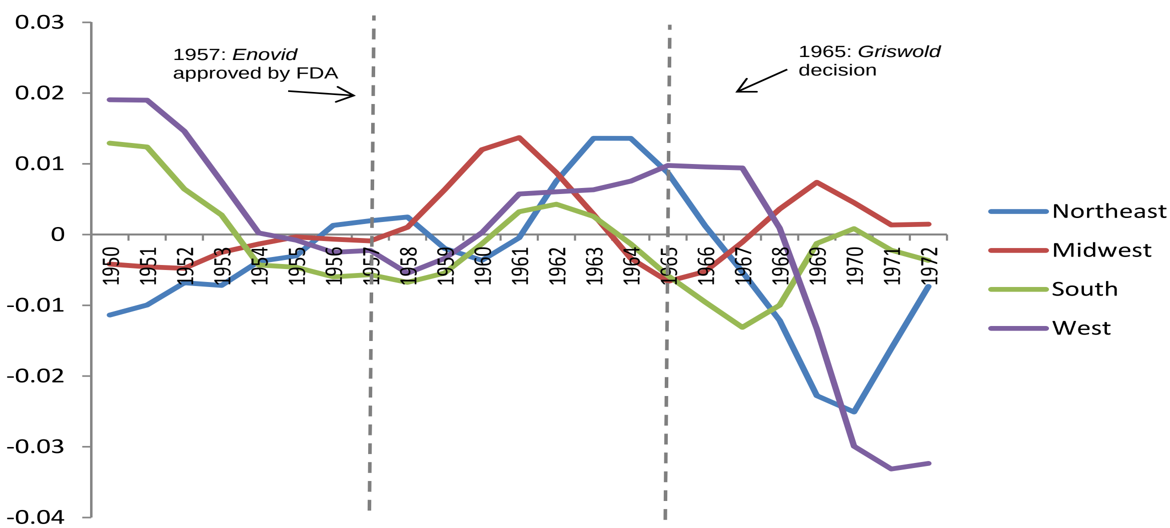
* See note in Figure 3.

**Figure 6:** Growth of Total Births, by Region: States with Sales Bans - States without Sales Bans



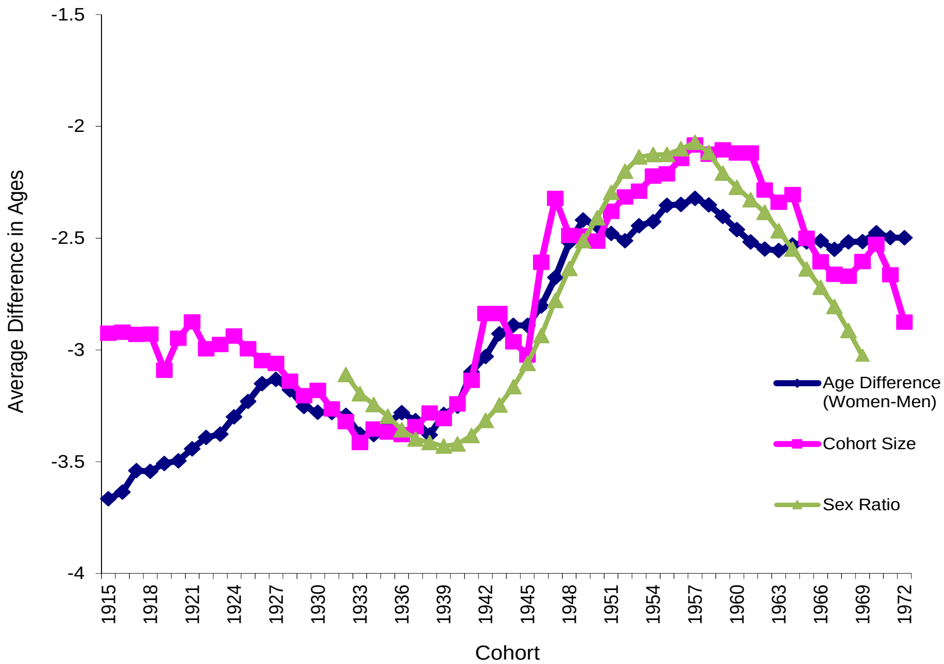

Notes: Three-year moving averages. Sources: National Vital Statistics. Bailey (2010).

**Figure 7:** Growth of Total Adult Population at Age 20, by Region: States with Sales Bans - States without Sales Bans




Notes: Three-year moving averages. Sources: NIH SEER population estimates. Bailey (2010).

**Figure 8:** Age Difference Between Spouses by Cohort, Whites




Sources: Vital Statistics of the United States; IPUMS CPS, 1962-2011; IPUMS USA, 1950-1960.

**Figure 9:** Observed and Simulated Never Married Rates for Women