Cohort Size and the Marriage Market:
Explaining Nearly a Century of Changes in U.S. Marriage Rates∗†‡§

Mary Ann Bronson‡ Maurizio Mazzocco§
Current Draft, June, 2016

Abstract
We document a strong and negative relationship between changes in cohort size and marriage rates for both women and men. We provide the most convincing evidence on this relationship by using variation in cohort size due to differences across states in sales bans on oral contraceptives. This empirical pattern is puzzling if interpreted using insights derived from the previous literature and standard matching models which indicate that an increase in cohort size should reduce the marriage rate of women, but increase the marriage rate of men. We then investigate the mechanism behind the negative relationship using a standard dynamic search model of the marriage market. We first show that the model we consider is rejected by the data for the same reason the standard matching model is rejected: it predicts that a rise in cohort size should reduce the marriage rate of women, but increase the marriage rate of men. We then develop two variations of the search model and show that they are both able to generate the observed relationship between cohort size and marriage rates. Lastly, we derive a testable implication for the two models based on the relationship between cohort size and divorce rates and provide evidence that only one model is not rejected by the data.

1 Introduction
Since the early 1970s, both economists and sociologists have analyzed the effect of demographic changes, such as population growth or migration, on marriage decisions. The idea that follows from this long-standing literature is simple. Whenever men are in low supply

∗We are grateful to Ran Abramitzky, Martha Bailey, Guillermo Beylis, Pierre-André Chiappori, Edgar Cortes, James Heckman, Joe Hotz, Caroline Hoxby, Kenneth Mirkin, Anne Pebley, Luigi Pistaferri, Dmitry Plotnikov, Michele Tertilt, Nico Voigtländner, 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.
†First Draft, February, 2012
‡Georgetown University, Dept. of Economics, Washington, D.C. Email: mary.ann.bronson@gmail.com.
§UCLA, Department of Economics, Bunche Hall, Los Angeles, CA. Email: mmazzocc@econ.ucla.edu.
1Prominent examples include Aker (1967), Becker (1973), Schoen (1983), Bergstrom and Lam (1989) and, more recently, Angrist (2002), Seitz (2009), Abramitzky, Delavande, and Vasconcelos (2011), and Knowles and Vandenbroucke (2016).
relative to women in the marriage market, women’s marriage rates decrease and men’s marriage rates increase. When women are in low supply, the opposite is true. This relationship is commonly rationalized using matching models.

Changes in cohort size, generated for instance by baby booms or busts, represent one of the most important sources of variation in the supply of women relative to men. Since women marry on average older men and women and men have similar cohort size at birth, a rise in cohort size will make older men a scarce resource and lead to a surplus of women relative to men in the marriage market. Using the insight developed in the previous literature, the consequence of a rise in cohort size should therefore be a decline in the marriage rate of women and an increase in the marriage rate of men.

This paper makes two main contributions linked to the relationship between changes in cohort size and marriage rates. First, we document that there is a systematic negative relationship between cohort size growth and marriage rates for women, which is consistent with the findings of previous papers. But, contrary to the insight provided by the existing literature and the predictions of a standard matching model, we document the same systematic, negative relationship between changes in cohort size and marriage rates for men. This puzzling result has been overlooked in previous papers, which have focused primarily on women. Our paper document these findings in two steps.

We first provide evidence on the existence of a strong and negative relationship between cohort size and marriage rates for both women and men using time-series variation and cross-sectional, state-level variation in cohort size and marriage rates. We find that, on average, a 10 percent increase in cohort size is associated with a 0.5 to 1 percent decrease in the share ever married by 40. Those estimates imply that changes in cohort size explain around 50% of the variation in marriage rates since the 1930s, a very large share. For comparison, important existing explanations for changes in marriage rates that have been tested empirically, like early access to “the Pill” in the 1970s, explain roughly 15% of the changes in marriage rates during the period for which they are most relevant (Goldin and Katz (2002)).

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 for women as well as for men, twenty to thirty years later. The exogenous variation, therefore, gives results that are consistent with the time-series and cross-state variation.
The second contribution of the paper is to investigate theoretically and empirically the potential mechanisms behind the negative relationship between changes in cohort size and marriage rates. To achieve that goal, we develop a dynamic search model of the marriage market. We use a search model instead of the standard matching model, because the latter is rejected by the data as it predicts that changes in cohort size should have opposite effects on the marriage rates of men and women. In addition, the search model enables us to easily capture the dynamic nature of the marriage market which, we believe, is an essential part of the mechanism behind the empirical results documented in this paper.

We use the model to show three main results. First, we show that a simple dynamic search model of the marriage market is rejected for the same reason the matching model is rejected: in the search model an increase in cohort size always increases the marriage rate of men. The matching model and the search model we develop have one common feature: changes in cohort size affect marriage rates exclusively through the relative number of women in the marriage market and the corresponding probabilities of meeting a spouse. The rejection of these two models indicates that the relationship between cohort size and marriage rates cannot be simply explained by changes in meeting probabilities.

Using this insight, we then develop two variations of the search model and show that they are both able to generate the empirical results documented in the first part of the paper. In the first version, a man can undertake a pre-marital investment that increases his probability of meeting a potential spouse and, if they marry, their marital surplus. In this alternative formulation, men are more likely to invest when cohort size is low because, in this case, the probability with which they meet a potential spouse is generally below the optimal level for a larger fraction of them. In the second version, the model is modified to allow the value of being single to increase with cohort size. The economic idea behind this new feature of the model is that it is more enjoyable to be single when cohort size is large because there are more individuals of the same age with whom to perform leisure activities.

The two versions of the search model have one common attribute. In both of them, an increase in cohort size has an additional negative effect on the marriage rate of women and men. We show that, if this additional effect is sufficiently strong, the search model can generate the observed pattern that a rise in cohort size generates a decline in the marriage rate of men. But, interestingly, the two alternative models generate this additional effect in different ways. The first version does this by increasing the value of marriage when cohort size declines, whereas the second version achieves this by increasing the value of being single when cohort size rises. This difference generates a testable implication based on the relationship between cohort size and divorce rates. The first version of the search model predicts that an increase in cohort size should generate an increase in divorce rates. The second version has the opposite prediction and generates a negative relationship between those two variables. As a final result, we use this implication to test one model against the
other. We document that, in the data, there is a strong and positive relationship between cohort size and divorce rates. We can therefore reject the model in which the value of being single increases with cohort size in favor of the investment model. This result suggests that the negative relationship between cohort size and marriage rates can be explained by changes in cohort size that simultaneously affect the relative supply of women in the marriage market and the marital surplus.

We conclude this section with one remark. The variable cohort size is not exogenous. The cohort size of individuals of marriageable age today was affected, twenty or thirty years earlier, by variables such as new fertility technologies, technological progress, child care availability, supply of housing, and changes in labor supply decisions of women. However, this fact does not diminish the importance of our results for social scientists and policy makers. It indicates that to understand the dynamics of the marriage market one has to study the evolution of cohort size and of the variables that have an effect on it. Finally, our results show that changes in cohort size are an excellent predictor of future trends in marriage and household formation, which in turn have first-order effects on aggregate labor supply, savings, and other macroeconomic variables of interest (Doepke and Tertilt (2016)).

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 and test the dynamic search model. Section 6 concludes.

2 Existing Explanations

In this section, we describe the papers that are most related to our work.

Akers (1967) and Schoen (1983) analyze the effect of cohort size on marriage rates using demographic models of the marriage market and data simulated from their models. In those models, changes in cohort size modify the supply of women relative to men in the marriage market, which in turn changes the marriage rate. The results of those papers suggest that an increase in cohort size should reduce the marriage rate of women and increase the marriage rate of men. Since those findings are the outcome of the developed models and of the corresponding assumptions, they should be interpreted as theoretical insights on the possible effect of changes in cohort size on marriage decisions. Our paper indicates that those models are rejected by the U.S. data since increases in cohort size reduce the marriage rate for both men and women.

There is an extended literature related to the previous two papers that makes use of changes in the sex-ratio, computed as the number of men divided by the number of women in the marriage market, to understand changes in marriage rates. The existence of a relationship between the sex ratio and marriage rates was formally postulated using a two-sided matching model in Becker (1973)’s seminal work. Since then, almost all papers studying
the marriage market have relied on matching models. An important contribution of Becker (1973) is to formally derive the testable implication that was informally discussed in previous work: a reduction in sex ratio should reduce women’s marriage rates and increase men’s marriage rates. Several papers have attempted to test the relationship between sex ratios and marriage rates. Angrist (2002), Seitz (2009), Abramitzky, Delavande, and Vasconcelos (2011), and Knowles and Vandenbroucke (2016) are examples of papers in that literature.²

In our paper, we do not explicitly analyze the link between sex ratios and marriage rates. But, as argued in Akers (1967), there is a relationship between changes in cohort size and changes in sex ratio under the standard and realistic assumption that women marry on average older men. Under this assumption, since men and women have similar cohort sizes at birth, the marriage market is composed of a similar number of younger women and men, but the number of older men is larger than the number of older women. In this context, an increase in cohort size increases the number of younger men and women relative to the number of older individuals. As a result, the rise in cohort size will have the effect of reducing the sex ratio. The matching model has therefore the following testable implication for changes in cohort size: an increase in cohort size should reduce the marriage rate of women, but increase the marriage rate of men. An important contribution of our paper is to provide evidence that this testable implication and, hence, the standard matching model is rejected if one considers U.S. data on cohort size and marriage rates over the past century.

This finding does not imply that a change in sex ratio can never generate a marriage market outcome that is consistent with the matching model. For instance, Abramitzky, Delavande, and Vasconcelos (2011) consider variation in the sex ratio due to World War I casualties in France and find that a smaller sex ratio is associated with a lower marriage rate for women and a larger marriage rate for men, in line with Becker’s prediction. Our findings simply indicate that, if the change in sex ratio is generated by a change in cohort size, the standard matching model is not consistent with U.S. marriage rate data from the past 100 years.

There is one paper in the sex ratio literature whose results are consistent with ours. Angrist (2002) uses variation in immigration rates from different European countries to the U.S. at the beginning of the twentieth century to study the relationship between sex ratios and marriage rates. He exploits the fact that the majority of migrants were men and that marriages were often formed between individuals belonging to the same ethnicity. Consistent with the patterns documented here, he finds that ethnicities with lower sex ratios experienced lower marriage rates for women as well as men. In our paper, we show that the positive relationship between sex ratio and marriage rates of women and men applies

²Several papers have studied the relationship between sex ratios and other economic variables, including rates of single motherhood (Neal (2004)), labor force participation of women (Grossbard-Shechtman (1984)), and birth rates (Bitter and Schmidt (2011)). Bergstrom and Lam (1989) calibrate a marriage matching model following Becker (1973) to study the relationship between cohort size, sex ratios, and age differences between spouses when marriage rates are held constant.
to a century of data and not only to the specific period considered in Angrist (2002), if the changes in sex ratios are produced by variation in cohort size. In addition, we propose two models that are consistent with our findings and show that they have opposite implications for the relationship between cohort size and divorce rates. Lastly, we use the derived implications to test one model against the other, therefore formally providing a mechanism that can explain the variation in marriage rates observed in the data.

There is one alternative explanation for the link between cohort size and marriage rates that has commonality with the one we propose: the Easterlin hypothesis. Easterlin (1987) argues that the relative size of a cohort can explain many variables that determine the economic and social outcomes of that cohort: earnings and unemployment rates, college enrollment rates, divorce, fertility, crime, suicide rates, and marriage. Easterlin’s explanation is composed of two parts. First, the distribution of income of a cohort is affected by its size, with larger cohorts having worse economic outcomes. Second, when income of a cohort is above its aspiration level, the individuals in that cohort will be optimistic and therefore will have better economic and social outcomes. Researchers who have attempted to test the general idea behind Easterlin’s hypothesis have found mixed results (Pampel and Peters (1995)).

To conclude, a number of other factors are likely to contribute to explaining changes in U.S. marriage rates over time. Technological changes that affect the benefits of being married relative to being single, such as the introduction of “the Pill” (Goldin and Katz (2002)), the legalization of abortion (Akerlof, Yellen, and Katz (1996)), and the decline in the price of household appliances (Greenwood and Guner (2009)), can help explain the period of rapidly falling marriage rates starting around the mid-sixties. Changes in men’s marriageability and labor market opportunities over time (Wilson (1987)) and in policies, such as welfare aid, may affect women’s desire to marry, though tests of these theories have yielded mixed results as discussed in an excellent review by Ellwood and Crane (1990). Finally, the rise in incarceration associated with the War on Drugs, which particularly affected black men, may help explain falling marriage rates especially among African Americans after 1980 (Charles and Luoh (2010)).

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 section Data Rising income and employment opportunities for women may also affect their desire to marry, though 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.
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 B1, 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 net migration 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 four parts. We first provide empirical evidence on the relationship between cohort size and marriage rates using longitudinal variation. We then describe findings obtained using cross-state variation. In the third subsection, we discuss endogeneity issues that may affect the longitudinal and cross-state variation. Lastly, 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.

When analyzing changes in marriage rates over time, existing studies have typically employed 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. In Bronson and Mazzocco (2016), we provide evidence that these two measures are not well suited to study the evolution of the share ever married for the following reasons. The first measure confounds changes in the numerator, new marriages, with changes in the denominator, population. If the population undergoes any substantive
growth or decline due, for instance, to changes in fertility or migration patterns, one will generally draw the wrong inference. The second 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. In the same paper, we also provide evidence that a measure that is not affected by those limitations is the share of individuals ever married by a given age in a given cohort, as long as the age-cutoff is sufficiently high. For this reason, in the rest of the paper we will only use this cohort-based measure.

4.1 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 1, 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. Panels A and B describe these variables for the white and black populations, respectively. We plot the share never married because visually it is easier to detect a positive correlation between the two variables. Figure 1 contains one main finding. To describe it, we initially focus on cohorts in Panel A born before 1960. For those cohorts, there is a strong positive correlation between cohort size and the share never married for both women and men. 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, when cohabitation started 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 2 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. Interestingly, once cohabiting households are accounted for, the relationship between cohort size and household formation for cohorts born in the 1960s and
1970s resembles 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 for the male and female population. Increasing cohort sizes in the second part of the 1970s are associated with a rise in the share never married and not cohabiting. Finally, Figure 2 shows that the strong positive correlation between cohort size and share never married characterizes both the white and the black populations.

In Figures 1 and 2, 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 3 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.

In Table 1, we record the average response of marriage rates to changes in cohort size at different age cutoffs. For ease of exposition, in the rest of the paper we will consider the effect of cohort size on the share ever married instead of the share never married. Each coefficient in the table is the outcome of a separate regression of the log share ever married or currently cohabiting on log cohort size. There are three main results in Table 1. First, the recorded elasticities are highest at 30, and gradually decrease with age, for both sexes and both races. This finding suggests that an increase in cohort size is associated with two effects: (i) a decrease in the eventual share ever married or cohabiting; and (ii) an increase in the age at first marriage. It also indicates that using 40 as an age cut-off better isolates the relationship between variation in cohort size and variation in marriage rates from changes in age at first marriage. The second result is that the effect of cohort size is quantitatively large. 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 0.85.4

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 subsection we will try to understand whether cohort size

---

4Note 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.
can explain the short-run, medium-run, or long run changes in marriage rates. Note that
the regressions in Table 1 document only an average relationship between changes in our
two main variables. To analyze whether changes in cohort size have to cumulate for several
years to generate a change 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 size for the cohort born in period \( t \)
for the \( n \)-year difference is constructed by adding up cohort size from \( t - n + 1 \) to \( t \).\(^5\) 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 esti-
mates for the 1-year differences indicate that a 1-year change in cohort size is not sufficient
to trigger a change in marriage rates. All the coefficients for the 1-year difference are statisti-
cally equal to zero, with the exception of the coefficient for women by age 30 and for men
by age 40.\(^6\) The effects are only marginally larger when we employ 2-year differences.

The negative effect of cohort size on marriage rates is much stronger when we study the
medium-term effect. Already at four-year differences, all the coefficients are large in size,
negative, and 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 for both men and women. 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 changes in cohort size and marriage rates using
cross-state variation. Because cohabitation has become a close substitute for marriage since

\(^5\)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.

\(^6\)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, which happen to coincide
with increases in share married for those two cohorts. If one drops the observations that characterize the
period around World War II, the coefficient becomes zero.
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 shorthand “ever married” we refer to those ever married or currently cohabiting.

4.2 Change in Marriage Rates Across States

In this subsection we provide evidence on the relationship between cohort size and household formation rates by using variation across states. If changes in cohort size generate variation in marriage rates, we should observe such an effect not just across time but also across geography. Specifically, states with larger increases in cohort sizes should experience larger drops in marriage rates.

When using cross-state variation, one additional issue we must address is that changes in cohort size at 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 cohort size at the time of marriage for each state and cohort. 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. 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 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 Table 3, the first stage results suggest that cohort size at birth explains a large fraction of cohort size at marriage age. They also indicate that the coefficient on log cohort size at birth, estimated to be 0.44, is well below 1, suggesting that cross-state migration significantly affects 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 sizes of the IV coefficients imply large effects that can explain a significant share of the changes in marriage rates. For example, in the two decades corresponding to the 1940 and 1950 cohorts, and the 1950 and 1960 cohorts, cohort size increased by substantial amount, 42% and 17%, respectively. Our estimates imply that this corresponded to a 3.4 and 3.8 percentage point decrease in the share of women and men ever married by 30 from 1940 to 1950, and a 1.5 and 1.4 percentage point decrease from 1950 to 1960. This is equivalent, on average, to about 55% and 39% of the actual changes in share ever married by 30 observed for women and men over this time period. 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.

4.3 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 for both women and men observed when longitudinal variation was used. However, there are reasons that prevent a causal interpretation of the relationship. While using number of births as an instrumental 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 increased births in 1950 relative to 1940 as well as a greater number of marriageable men in 1980 relative to 1970. Alternatively, states may differ in their religiosity or preferences for forming a family. 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.

4.4 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 birth control pill, and cross-state variation in anti-obscenity laws, which limited the use of contraception, to generate exogenous variation in the number of births and therefore cohort size.

In 1873, the U.S. Congress enacted the Comstock Act which had two main objectives. The direct goal 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 goal 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.

Because the specific language of these statutes varied significantly across states, the 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 consists of the seventeen states that explicitly banned the sale, advertisement, and distribution of information of any product for the prevention of
conception. The second group consists of seven states that had the same ban 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 related to birth-control methods. The third category is comprised of six states that explicitly only banned the advertisement and distribution of information about contraceptives, but did not outlaw their sale. The final group is composed of eighteen states that approved a law banning the sale, advertisement, and dissemination of information of obscene products, without explicitly classifying the prevention of conception as obscene. In our analysis, we refer to states in the first two groups as having a sales ban on contraceptives. 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 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 total births, and therefore cohort size.

We will start by giving some 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 two variables: cohort size at birth and cohort size at marriage age. We follow Bailey (2010) and present the results separately for the four Census regions.
In Figure 4, Panel A, 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. Two features generate the hump shape in the difference in growth of number of births characterizing all census regions. First, in all regions, after the introduction of the pill, states with the ban experienced larger growth in cohort size at birth. In the South, states with the ban were already increasing relative to states without a ban before 1957, but the process was expedited by the introduction of the pill. Second, 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.

In Figure 4, Panel B, 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 in Panel A, with the differences in growth in cohort size following the same familiar hump shape between 1957 and 1965. Thus, the differences in cohort size at birth driven by the Comstock laws and the introduction of the Pill persisted to generate differences in cohort size at marriageable age.

We now formally use the introduction of the pill interacted with the sales bans as an instrument for cohort size at marriageable age. To that end, we construct two dummy variables: the first one, \( \text{ban}_s \), is equal to one for all cohorts from state \( s \) that adopted a sales ban on contraceptives and zero otherwise; the second dummy variable, \( \text{ban} \times \text{pill}_{c,s} \), 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. In the first stage, we then 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.

\[
\log \frac{y_{c,s}}{y_{c-n,s}} = \alpha + \beta_1 \text{ban}_s + \beta_2 \text{ban} \times \text{pill}_{c,s} + \sum_{c,r} \pi_{c,r} + X' \gamma + \varepsilon_{c,s},
\]

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 \( \text{pill}_{c,s} \), and an indicator equal to 1 if the state enacted an advertising ban on contraception.

The results of the first stage are presented in Table 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. 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. Consistent with Figure 4, Panel B, the effect is larger as we go from a 1-year difference to a 5-year difference, while the coefficients for the 5-year and 7-year differences are similar in size. Our estimates on the 5-year and 7-year differences indicate a 4.0 – 4.1% higher growth in cohort size in states with the sales ban between 1957-1965.
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 \( \textit{ban}_s \) 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.

\[
\log \frac{\text{mar}_{c,s}}{\text{mar}_{c-n,s}} = \beta_0 + \beta_1 \log \frac{\text{size}_{c,s}}{\text{size}_{c-n,s}} + \sum_{c,r} \pi_{c,r} + X'\gamma + \varepsilon_{c,s},
\]

except that now we instrument cohort size growth with \( \text{ban}_s \) and \( \text{ban} \ast \text{pill}_{c,s} \). 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 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 for both women and men. 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. Note, however, that we would expect the IV coefficients to be negative and larger in magnitude given the discussion in section 4.3 on potential endogeneity concerns, since the 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, a strong predictor of parental socioeconomic status. We employ the same specification used in the first stage of the IV estimation except that, to make our results comparable with Ananat and Hungerman (2012)’s findings, we use levels instead of differences for the following 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 all women who gave birth in the period of interest, the majority of whom are married, 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.7

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, suggesting that the young women in their sample period who reduced their unwanted pregnancies using the Pill may have been positively selected. Bailey (2013) tests for such selection during our relevant sample period in the 1960s using data from the Integrated Fertility Survey Series (IFSS). The IFSS includes socioeconomic measures such as race and education, as well information about the wantedness and timing of pregnancies and births of female respondents. Bailey documents that Pill usage in the early 1960s was concentrated among women in married

7A 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. Using the same exogenous variation as we do, but a slightly different specification, Bailey (2013) similarly finds no evidence that the birthweight of infants born in the 1960s changed differentially in states where selling the Pill was legal from states where it was not.
households, as expected. However, Bailey does not find evidence that reductions in unwanted births were higher for highly-educated women. Finally, we want to note that an increase in the share of children born in poor families 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.\footnote{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.}

To conclude, we have provided evidence on two main results. First, there is a causal negative relationship between cohort size and marriage rates of women. Second, there is a similar causal negative link between cohort size and marriage rates of men. The rest of the paper proposes and tests potential mechanisms that can explain our empirical results.

5 A Dynamic Search Model of the Marriage Market

In this section, we develop and test a model that has the potential of generating the two empirical patterns observed in the data. These patterns are puzzling because they contradict the prediction of a standard matching model, the model most commonly used to analyze marriage markets.\footnote{Some examples of papers studying marriage markets using matching models include Gale and Shapley (1962), Becker (1973), Becker (1974), Mortensen (1988), Bergstrom and Lam (1989), Bergstrom and Bagnoli (1993), Angrist (2002), Peters and Siow (2002), Choo and Siow (2006), Iyigun and Walsh (2007), Chiappori, Iyigun, and Weiss (2009), Hitsch (2010), and Abramitzky, Delavande, and Vasconcelos (2011).} As discussed in Section II, the matching model predicts that an increase in cohort size should raise the marriage rates of men, rather than lower them, as we observe in the data.

Since the matching model is rejected, in this section we construct a dynamic search model of the marriage market and test whether it can explain our main empirical findings. A priori, this model has the potential of producing the patterns observed in the data for reasons that will be discussed in the next subsection. In addition, the search model enables us to easily capture the dynamic nature of the marriage market which, we believe, is an essential part of the mechanism behind the empirical results documented in this paper.

5.1 Characterization of the Model

We start by developing the simplest possible version of a dynamic search model. We then evaluate whether it can generate the empirical patterns observed in the data.

Consider an economy populated by $T + 1$ overlapping generations of men and women. In each period $t$ a new generation, or cohort, is born and lives for $T + 1$ periods. $N_{0,t}^w$ and $N_{0,t}^m$ denote the size of the new generation of women and men born in period $t$. We assume
that women and men have the same cohort size $N_{0,t}$, which is a good approximation for the white population. In the Appendix we consider the more general case in which $N_{0,t}^w \neq N_{0,t}^m$.

Men and women can be either single or married. The within-period utility of being single is denoted by $\delta$, whereas the within-period utility of being married for the couple as a whole is denoted by their match quality $\eta$.

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 individuals then decide whether to marry with the objective of maximizing their lifetime utility. Once married, they make no further decision. The value of being single is constant across individuals and over time. The value of being married is drawn from a distribution $F(\eta)$ which does not depend on the age of the couple or time. 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 or until they divorce. Divorce can occur in each period with a probability $q$ which does not depend on the within-period utility, time, or age. In the next subsection, we will discuss the consequences of allowing the probability of divorce to depend on the quality of the marriage. A couple can freely divide the gains from marriage using a Nash-bargaining solution, with the parameter $\gamma \in [0, 1]$ determining how the marriage surplus is divided. Future utilities are discounted at the discount factor $\beta \leq 1$.

We can now introduce the main assumption of the model. We assume that single women meet a potential spouse with a positive probability only in their first period of life, while single men meet a possible partner 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 period of fertility. Second, an 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 this general concept.\footnote{This is not the first paper to use the differential fertility between women and men to develop a model of the marriage market. For instance, Siow (1998) uses a similar idea.}

Our main assumption has two implications. First, the marriage market is populated by younger women (age 0) and younger and older men (age 0 and 1). Second, women cannot marry after the first period and men cannot marry 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 solution of search models is generally provided in terms of reservation values. In our model, the relevant reservation value is the match quality $\eta$ at which a pair of potential spouses is indifferent between marrying or staying single. In our context, two types of couples can form: couples in which the woman is younger and the man is older; and couples
in which both the woman and the man are younger. Our model is therefore characterized by two reservation utilities.

In Appendix A.1, we show that a couple with an older man has the following reservation match quality:

$$\eta_{1,t} = 2\delta.$$  

The intuition behind this result is straightforward. A woman and an older man choose to marry if they draw a match quality $$\eta$$ greater than the sum of their values of staying single, $$2\delta$$. The reservation value of a couple in which the man is younger is slightly more complicated to derive because a younger man has the option value of waiting until next period and drawing a new potential spouse. In Appendix A.1, we show that the existence of this option value generates a reservation utility for this type of couple of the following form:

$$u_{0,t} = 2\delta + B\theta_{m,1,t+1},$$  

where $$\theta_{m,1,t+1}$$ is the probability that a man meets a woman when older, and $$B$$ depends only on the parameters of the model. The option value is included in the term $$B\theta_{m,1,t+1}$$, which measures the probability that a younger man will meet a woman when older multiplied by the share of the expected marital surplus he will receive times the probability he will choose to marry her.

We will now use the derived reservation utilities to solve for the steady state equilibrium in the marriage market. We will then evaluate how a change in cohort size from its steady state level impacts marriage rates. To solve for the steady state equilibrium, we have to derive the probability that a younger man meets a woman $$\theta_{0,t}$$ and the corresponding probability for an older man $$\theta_{m,1,t}$$. Let $$N_{i,a,t}$$ be the number of individuals of gender $$i$$, age $$a$$, in period $$t$$ who are present in the marriage market. Then, under our assumption that men and women have identical cohort size, $$\theta_{0,t}$$ and $$\theta_{m,1,t}$$ can be derived by noting that

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

The probability $$\theta_{0,t}$$ is the correct measure in our model of the sex-ratio, defined as the number of women divided by the number of men in the marriage market. Equation (3) shows that an increase in cohort size $$N_{0,t}$$ increases the relative supply of women in the marriage market, as argued earlier in the paper.

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

$$N_{1,t}^m = N_{0,t-1} \left(1 - \theta_{0,t-1}^m\right) + N_{0,t-1} \theta_{0,t-1}^m F\left(\eta_{0,t-1}\right).$$

(4)

In Appendix A.2 we show that, in steady state, equation (4) simplifies to

$$N_1^m = N_0 F\left(\eta_0\right)^{\frac{1}{2}}.$$

Substituting for $N_1^m$ in the equation defining the meeting probabilities $\theta_a^m$, we obtain

$$\theta_0^m = \theta_1^m = \frac{N_0}{N_0 + N_0 F\left(\eta_0\right)^{\frac{1}{2}}} = \frac{1}{1 + F\left(\eta_0\right)^{\frac{1}{2}}}.$$

To determine the reservation value of younger men in steady state, we can substitute for $\theta_1^m$ in the equation that determines the reservation value (2) and obtain

$$\eta_{ss} = 2\delta + B \frac{1}{1 + F\left(\eta_{ss}\right)^{\frac{1}{2}}}.$$

Note that $F(\eta)$ is monotonically increasing in $\eta$. As a consequence, there is a unique solution for $\eta_{ss}$ and hence a unique steady state equilibrium. Moreover, in the steady state equilibrium the reservation value is independent of cohort size $N_0$. The following Proposition summarizes the result.

**Proposition 1** In steady state, there is a unique reservation value for marriage $\eta_{ss}$ and hence a unique equilibrium. Moreover, the reservation utility does not depend on cohort size.

Now that we have characterized the steady state equilibrium we can study the effect of a change in cohort size first on the marriage rate of women and then on the marriage rate of men. We will focus on the case in which the shock is unexpected. Similar results apply if the shock is known with certainty.

Suppose the steady state economy is hit by a 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. The change in cohort size affects the marriage rates by changing the probability that an individual meets a potential spouse and, as a consequence, the reservation match quality of younger men. The following Proposition establishes that an increase in cohort size raises the reservation utility of younger men.

**Proposition 2** A positive and permanent shock to cohort size in period $\tau$ increases the reservation value $\eta_{0,\tau}$. A negative shock has the opposite effect.

**Proof.** See Appendix A.3. ■
The intuition behind this result is straightforward. With a permanent increase in cohort size, men are more likely to meet a woman when older. As a consequence, the option value for younger men of waiting until next period goes up and with it their reservation match quality.

Using Proposition 2 we can now study the effect of a cohort shock on the marriage rate of women and men. The following Proposition determines the effect for women.

**Proposition 3** A positive and permanent shock to cohort size in period $\tau$ reduces the fraction of cohort $\tau$ women who get married. A negative shock in period $\tau$ has the opposite effect.

**Proof.** See Appendix A.4. ■

To provide the insight behind this result, consider an increase in cohort size. After this event, older men become a scarce resource. This change has two effects. First, the fraction of women who marry mechanically declines because women are now less likely to meet older men. Second, younger men become more selective because they will have a larger group of women to choose from when they are older. As a consequence of this second effect, 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 above can explain the negative relationship observed in the data between cohort size and marriage rates for women.

The following Proposition establishes that the search model is not as good at generating the observed patterns for the marriage rates of men.

**Proposition 4** A positive and permanent shock to cohort size in period $\tau$ increases the fraction of cohort $\tau$ men who get married. A negative shock in period $\tau$ has the opposite effect.

**Proof.** See Appendix A.5. ■

Proposition 4 contains a negative result. Since our simple dynamic search model generates a positive relationship between cohort size and marriage rates of men, it is rejected by the data. We want to point out, however, that without a formal proof this result is not obvious. To understand why, observe that an increase in cohort size has two different effects on men. First, the probability that younger and older men in a cohort meet a woman increases. Second, younger men become more selective because they will have more women to choose from when they are older. The first effect go against our empirical findings since they imply an increase in the marriage rate of men when cohort size increases. But the second effect is in our favor and generates a decline in the marriage rate of men. Proposition 4 establishes that the first two effects always dominate, therefore rejecting our simple search model.

Our results clearly reject the simple search model developed in this section. There are, however, two modified versions that can generate the two main patterns documented in
the data. The next subsection develops them, derives a testable implication based on the relationship between cohort size and divorce rates, and provides evidence that only one of the two versions is consistent with that implication.

5.2 Two Augmented Dynamic Search Models of the Marriage Market

An important feature of the search model considered in the previous subsection is that changes in cohort size affect marriage decisions only through the probability of meeting a potential spouse. Its rejection indicates that in the data changes in cohort size affect household formation in ways that go beyond their impact on meeting probabilities.

There are two approaches that can be used to enrich the search model and reconcile it with our empirical findings. The first possibility is to allow men to undertake a costly investment that increases their probability of meeting a woman and, in case they choose to marry, the corresponding utility from marriage. The economic insight underlying this model is that men can choose to work longer hours or higher-paid jobs with fewer amenities to increase the probability of meeting a potential spouse and their marital surplus (Peters and Siow (2002), Iyigun and Walsh (2007), and Chiappori, Iyigun, and Weiss (2009)). In this model, the fraction of men who select to invest is higher when cohort size is relatively low, since in this case the probability of meeting a woman will generally be lower than optimal for a higher fraction of them. We will denote this model with the term investment-model. The alternative way of reconciling the search model with the data is to introduce a positive correlation between the value of being single \( \delta \) and cohort size. The economic idea behind the positive correlation is straightforward: when cohort size increases it is more enjoyable to be single because there are more people of the same age with whom to perform different types of leisure activities. We will refer to this model as the \( \delta \)-model.

In the two models, changes in cohort size generate an additional negative effect on marriage rates. Both have therefore the potential of explaining the two main empirical findings of our paper. The additional effect strengthens the negative relationship between cohort size and marriage rates generated by the simple search model for women. It also introduces a new negative effect for men which, if strong enough, can outweigh the effects that in the search model produce the positive relationship between cohort size and marriage rates of men. The last consideration implies that the two models can generate the observed patterns only for some parameter choices. If the additional negative effect is not strong enough, the two models will produce the same positive relationship between cohort size and marriage rates of men that characterizes the simple search model. It is therefore not productive to search for a general proof showing that our models can produce the observed patterns for men. Instead, we will calibrate the models and evaluate whether they can match our empirical findings for a realistic set of parameters.

An interesting feature of the two models is that they generate the additional negative
effect in different ways. The investment-model achieves this by increasing the value of marriage when cohort size is low, whereas the \(\delta\)-model does this by increasing the value of the outside option when cohort size is high. As a result, if the probability of divorce declines with the match quality of the marriage, our two models generate relationships between cohort size and divorce rates of opposite sign. The investment-model predicts a positive relationship between cohort size and divorce rates. When cohort size decreases, men invest more and, by doing so, they increase the match quality of their marriage. The probability of divorce will therefore decline, generating a positive relationship between cohort size and divorce rates. The \(\delta\)-model has the opposite prediction. When cohort size declines, the value of being single becomes lower. As a consequence, the average match quality of formed marriages drops, which increases the probability that a marriage will end in divorce.

We will use this result to test the \(\delta\)-model against the investment-model. To implement this test, we need to slightly modify the way divorce is modeled and allow its probability to depend on match quality. This change should have no effect on the rejection of the simple search model based on Proposition 4. When the probability of divorce declines with the value of marriage, younger men become more selective because the value of marrying a woman with higher match quality increases. However, this increase in reservation utilities of younger men will be similar for all cohort sizes, implying that the marriage rate of men will still rise with cohort size. Figure A1 in the Appendix confirms this insight using simulations obtained from the original search model when the probability of divorce is a decreasing function of match quality.

We will now describe how we modify the search model to obtain the investment-model and \(\delta\)-model. In the investment-model, men can make a decision when young to undertake an investment that increases their attractiveness in the marriage market. If they choose to invest, they pay a cost \(c_i\), which is individual-specific. The investment has two benefits. First, men who invest meet a woman with probability \(\theta_h\), whereas men who do not invest meet a woman with probability \(\theta_l\), with \(\theta_h > \theta_l\) in all periods. The economic insight for this feature of the model is that it is easier for more attractive men to draw the attention of women. Second, a man who invests experiences a higher level of match quality if he gets married. Specifically, if he and his spouse drew a match quality \(\eta'\), they enjoy a within-period value of marriage \(\eta = \eta' + K\), where \(K\) is a positive constant. The idea behind this modeling choice is that men who invest have higher income and wealth and can therefore afford to buy more public goods when married.

In this version of the model, whenever cohort size increases and the probability of meeting a woman is high, men have less incentive to invest and investment rates are low. Since foregoing a pre-marital investment makes marriage between two potential partners less likely, the share of men who marry decreases, generating a negative relationship between changes in cohort size and marriage rates.
It is useful to discuss the intuition for why in this model investment has two benefits instead of just one. If investment affects only match quality but not meeting probabilities, men would be more, rather than less, likely to invest when cohort size increases, since in those periods they would have a higher probability of meeting a woman and, hence, a higher probability of realizing a return on their pre-marital investment. This would generate a positive rather than a negative relationship between changes in cohort size and marriage rates. Alternatively, if investment affects meeting probabilities but not match quality, its only effect is to switch women from men who do not invest to men who invest without changing the marriage rate.

We calibrate the new parameters of the investment-model as follows. First, we assume that the individual cost of investing \( c_i \) takes the following functional form:

\[
c_i = \mu_0 + \mu_1(1 + x_i)^{\mu_2},
\]

where \( x_i \) is the type of the individual and is drawn from a uniform distribution defined on the interval \([0, 1]\). The cost parameters \( \mu_0, \mu_1, \) and \( \mu_2 \) influence the distribution across cohorts of the share of men who invest. While we cannot observe such a measure directly, we use as a proxy the share of men working full-time and full-year or studying at age 20 and match its distribution over cohorts. The power coefficient \( \mu_2 \) is necessary to match how the share of men who invest changes with cohort size. Without this parameter, or if this parameters is too low, the model generates overly large swings in investment due to the following amplification effect. When cohort size declines and the probability of meeting a woman falls, more men choose to invest. The investment further lowers the probability of meeting a woman for the remaining men who did not invest, increasing their incentives to undertake the investment.

To calibrate the investment parameter \( K \), we rely on the observation that \( K \) affects the share of younger men who choose to marry since, for a larger \( K \), by not marrying they forgo a larger marital surplus in the current period. Using this idea, we calibrate \( K \) by matching the share of men who marry when younger, which we assume to be by age 30. Lastly, in the model we assume that \( \theta_h \) is equal to a constant, while \( \theta_l \) is endogenous and equal to the number of younger women who did not meet a man who invested divided by the number of men who chose not to invest. In the data we do not observe the probability of meeting a woman for a man who invests. But we know that, to match the marriage patterns in the data, \( \theta_h \) must be sufficiently high so that it is always larger than \( \theta_l \). If not, men would choose either not to invest or to invest in periods when cohort size is increasing and the probability of meeting a woman is high. To satisfy that restriction, we have experimented with values of \( \theta_h \) that are between 0.7 and 1, obtaining similar results. The simulations presented in this paper have been generated using \( \theta_h \) equal to 0.85.

Deriving the \( \delta \)-model is more straightforward, since we implement one simple change.
We allow the value of being single to be an increasing function of cohort size, where the function takes the following form:

\[ \delta_t = \alpha_0 + \alpha_1 N_{0,t}. \]

The linear functional form has been chosen for simplicity. We have also experimented with alternative functional forms that are increasing in cohort size with similar results. The new parameters \( \alpha_0 \) and \( \alpha_1 \) affect the share of individuals who choose to stay single. All else equal, a higher value of \( \alpha_0 \) increases the share of people who choose to remain single independently of cohort size, whereas a higher value of \( \alpha_1 \) raises proportionally more the share of never married individuals born in larger cohorts. Using this idea, we calibrate these parameters by matching the minimum and the average share never married in the data, computed over the cohorts observed in our sample period.

In both versions of the search model, we set the parameter \( \gamma \) governing how the marital surplus is allocated to 0.5, and assume that the distribution of match quality \( F(\eta) \) is uniform in the interval \([0, 1]\). Finally, in both models, we allow the probability of divorce \( q \) to decline with match quality \( \eta \) using the following linear functional form:

\[ q(\eta) = \rho_0 + \rho_1 \eta, \]

where \( \rho_1 \leq 0 \). The parameter \( \rho_0 \) affects the share of households who choose to divorce independently of match quality. The parameter \( \rho_1 \) influences how the share of divorces changes with match quality. In the simulations, we restrict \( \rho_0 \) and \( \rho_1 \) so that \( 0 \leq q(\eta) \leq 1 \) for all values of \( \eta \), and calibrate the two parameters by matching the minimum and the average share of individuals ever divorced, computed over the cohorts observed in our sample period.

To construct these data moments, we need to observe current divorce status, as well as the number of previous marriages for all currently married individuals. The latter variable is available in the American Community Survey (ACS), but only for the years 2008-2013. In those years, we construct the share ever divorced for all individuals between the ages of 50 and 75, i.e. individuals who are sufficiently old to have had a chance to experience a divorce. This corresponds to cohorts born between 1933 and 1963.

Figure 5 documents the performance of the two versions of the search model. Like the simple search model, the investment-model and \( \delta \)-model replicate the negative relationship between cohort size and marriage rates of women. But, unlike the basic search model, the two variations can match the negative relationship for men observed in the data. They are therefore both consistent with our empirical findings.

Since both proposed models can match well our empirical results, we now use the testable implication based on the relationship between cohort size and divorce rates to evaluate which framework is a better characterization of the data. At the beginning of this subsection we
argued that the investment-model predicts a positive relationship between cohort size and divorce rates, whereas the \( \delta \)-model predicts a negative relationship. Figure 6, Panel A, documents the share of individuals in the data ever divorced for cohorts born between 1933 and 1963. There is a strong and positive relationship between cohort size and the share ever divorced, for both women and men. Individuals born between 1933 and the first half of the fifties experienced increasing cohort size and rising divorce rates, whereas people born in the second half of the fifties and the first half of the sixties are characterized by declining cohort size and falling divorce rates. The evidence provided in Figure 6 and the previous discussion clearly imply that the \( \delta \)-model is rejected in favor of the investment-model. Panels B and C, which report the simulated share ever divorced for the investment-model and \( \delta \)-model, confirm this conclusion. The investment-model is the only framework able to generate the positive relationship between cohort size and share ever divorced. In that model, the share of women and men ever divorced peaks a bit early relative to the data, but otherwise matches well the observed patterns.

A striking feature of the data displayed in Figure 6A is that, when the time-series of cohort size starts to flatten, the share of women ever divorced crosses from below the share of men ever divorced. Interestingly, the investment-model generates the same type of crossing. To see why, recall that, all else equal, the marriages with the lowest probability of divorce in the model are those in which the man has made an investment. In the first half of the thirties, the fraction of men who invested dropped as cohort sizes began to increase, after having declined for over a decade. Correspondingly, both men’s and women’s divorce rates started to increase. However, women’s divorce rates remained lower than men’s, since a significant share of those women were married to older men, who came from smaller cohorts and had higher rates of investment. Over time, the share of women in such marriages dropped steadily. Accordingly, in Figure 6, Panel B, the difference between the probability of divorce of women and men gradually shrinks until the two variables become approximately equal for cohorts born around 1950. Cohorts born after 1950 display the opposite pattern. Men began investing again at increasing rates as they experienced first a flattening in cohort size growth and then a significant drop. Correspondingly, divorce rates of men began to fall. Women’s divorce rates did not begin to fall until later, since many of the women in those cohorts married older men from larger cohorts who had lower rates of investment. As a consequence, the share of women ever divorced stays above the share of men for the rest of the sample period.

Divorce rates start to fall earlier in our simulations than in the data. One possible explanation for the difference is that divorce laws changed significantly in the 1970s, when cohorts born in the 1950s were entering the marriage market. Since the model abstracts from this aspect of the data, it is to be expected that our simulations generate lower divorce rates for cohorts born after 1950. Nevertheless, it is remarkable that the model matches
the general patterns in divorce rates using only cohort size, without incorporating the legal changes to the marriage contract taking place during our sample period.

We end by providing two additional pieces of evidence in support of the explanation developed in this section. First, using Census and ACS data from 1940 to 2010, we document the evolution of a measure of the fraction of younger men who invest: the share of men either in college or working full-time, full-year at age 20. We use this measure because it represents a proxy of the fraction of younger men who are committed to accumulating financial resources and/or human capital. This measure is plotted in Figure 7 together with the fraction of younger men who invest generated by our model. Marriage considerations are only one of several factors that explain labor supply and education decisions. Nevertheless, the model can match reasonably well the variation observed in the data, including the increase in the share investing for cohorts born at the beginning of the century, the decline for the cohort born in 1950, the rise for cohorts born in 1960 and 1970, followed by a flattening and decline for cohorts born in the last part of our sample. Only for the 1940 cohort does our model make the incorrect prediction that the share investing declined. As mentioned above, it is to be expected that the model cannot explain all the variation in the measure of investment given the complexity of the data and the simplicity of the model.

The second piece of evidence in support of the explanation developed in this section concerns the relationship between cohort size and age differences between spouses. An implication of our search model is that the relationship should be negative. Since an increase in cohort size makes older men a scare resource, women marry on average younger men when cohort size rises. As a result, the average age difference between spouses should decline with an increase in cohort size. To test this prediction, in Figure 8 we report the evolution of the average age difference between spouses in the data and in the simulations, and the evolution of cohort size. With the exception of the first twelve cohorts, Figure 8 indicates that there is a tight negative relationship between age difference and cohort size in the data and that our model can replicate well that pattern. 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 in both the data and simulations. When cohort size drops, the age difference becomes more negative and therefore increases.

The results presented in this section may also help explain why some researchers have found that marriage rates of men decline when the relative supply of women drops (Abramitzky, Delavande, and Vasconcelos (2011)), whereas others (Angrist (2002)) provide evidence that marriage rates of men increase. According to our findings, a decline in relative supply of women will generate an increase in marriage rates of men if they have the incentive to undertake an investment that boosts their attractiveness. This is generally the case in a marriage market in which there are more men than women: men will undertake the investment to draw the attention of the limited number of women. Since the fertility
stage of women is shorter than for men and one of the main reasons for marriage is to have and raise children, in the U.S. marriage market men outnumbered women in most years. Similarly, in the marriage market studied by Angrist (2002) for different ethnicities with high immigration rates men generally exceeded the number of available women. In those instances, men have the right incentives to invest and one should expect a rising marriage rate for men when the relative supply of women drops, which is the finding in our paper and in the paper by Angrist (2002). By contrast, Abramitzky, Delavande, and Vasconcelos (2011) consider marriage markets in France after World War I where, due to the high mortality rate of men, in most regions women exceeded men (in some areas the sex ratio was as low as 864 men per 1,000 women). In that context, men have limited incentives to invest to attract women and their marriage rate should follow the pattern predicted by the standard matching model: it should decline with a reduction in the relative supply of women as observed in Abramitzky, Delavande, and Vasconcelos (2011).11

To summarize, the findings of this section indicate that a simple dynamic search model with investment can generate the marriage and divorce patterns observed in the data. This is not an easy task, since some of the patterns, such as the crossing in the divorce data, are not a priori easy to explain.

In principle, it should be possible to explain the same patterns by adding investment to the standard matching model. We have used the search model instead, because it is better suited to capture how changes in cohort size affect the relative supply of women in the marriage market and, hence, marriage and divorce decisions. Our results should, therefore, not be interpreted as a rejection of the matching model in favor of the search model. They should be understood as indicating that the effects of cohort size on marriage decisions are sufficiently complicated that a standard search or matching model cannot explain them. To rationalize the data an additional force must be added that increases the value of marriage when cohort size declines.

6 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 quasi-random variation in the adoption of the pill across states we provide evidence in support of the following two results. First, cohort size can explain on its own about 50% of the variation in U.S. marriage rates. Second, an increase in cohort size reduces marriage rates for both women and men and a decrease has the opposite effect.

11In our paper, we focus on investment by men because, as mentioned above, we consider marriage markets in the U.S. where men generally outnumbered women. Pre-marital investment by women is an aspect that should certainly be considered in markets in which women exceed men. Abramitzky, Delavande, and Vasconcelos (2011)’s findings suggest that women’s investment was not sufficiently large to change the prediction of the standard matching model in France after World War I.
We then develop and test a model that can generate the empirical patterns observed in the data. An important implication of our results is that the standard matching model is not consistent with our empirical findings, since it predicts that a rise in cohort size should increase the marriage rate of men. For this reason, we propose and test a dynamic search model of the marriage market. Using this model we document three results. First, a standard dynamic search model is rejected by the data for the same reason the standard matching model is rejected: it predicts that a rise in cohort size should always increase the marriage rate of men. We then show that two variations of the standard search model can generate all the patterns observed in the data. In the first variation, men can choose to undertake an investment that increases their probability of meeting a potential spouse and, if they marry, their marital surplus. In the second variation, the search model is modified by allowing the value of being single to be an increasing function of cohort size. An interesting feature of the two variations of the search model is that they imply a relationship between cohort size and divorce rates of opposite sign. Using this testable implication, we reject the second model in favor of the investment model.

Our results have important implications for policy analysis. Since the 1990s, politicians and policy makers have frequently considered and at times implemented policies to increase the fraction of married individuals with the intent of reducing the poverty rate. Several examples of such policies exist, such as “First Things First” in Tennessee, “Healthy Marriages Grand Rapids,” or the federal Temporary Assistance to Needy Families (TANF) program, which allows states to use a fraction of its funds to implement policies aimed at increasing the share of married individuals. Our findings indicate that these types of policies are likely to have only short-term effects, since in the medium and long run marriage rates are in large part driven by demographic changes that are difficult to control. Our findings also suggest that, given demographic changes, policies focused on increasing pre-marital investments of young individuals should be effective at increasing marriage rates.
References


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

<table>
<thead>
<tr>
<th></th>
<th>White Men</th>
<th>Black Men</th>
<th>White Women</th>
<th>Black Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ever married by age 30</td>
<td>-0.294***</td>
<td>-0.592***</td>
<td>-0.193***</td>
<td>-0.870***</td>
</tr>
<tr>
<td>R²</td>
<td>0.037</td>
<td>0.048</td>
<td>0.026</td>
<td>0.064</td>
</tr>
<tr>
<td>Ever married by age 35</td>
<td>-0.182***</td>
<td>-0.440***</td>
<td>-0.111***</td>
<td>-0.560***</td>
</tr>
<tr>
<td>R²</td>
<td>0.015</td>
<td>0.031</td>
<td>0.011</td>
<td>0.043</td>
</tr>
<tr>
<td>Ever married by age 40</td>
<td>-0.107***</td>
<td>-0.322***</td>
<td>-0.066***</td>
<td>-0.453***</td>
</tr>
<tr>
<td>R²</td>
<td>0.008</td>
<td>0.019</td>
<td>0.007</td>
<td>0.026</td>
</tr>
</tbody>
</table>

*** Significant at 1%. 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 USA 1960-1970.

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

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

* Significant at 10%. ** 5%. *** 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

<table>
<thead>
<tr>
<th>Dependent Variable: 10-Yr. Difference in Log Share Ever Married</th>
<th>OLS</th>
<th>IV (1)</th>
<th>IV (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Log Cohort Size</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men By Age 30</td>
<td>-0.041**</td>
<td>-0.123***</td>
<td>-0.168***</td>
</tr>
<tr>
<td><em>R</em>^2</td>
<td>0.807</td>
<td>0.791</td>
<td>0.819</td>
</tr>
<tr>
<td>Women By Age 30</td>
<td>-0.029***</td>
<td>-0.047**</td>
<td>-0.040</td>
</tr>
<tr>
<td><em>R</em>^2</td>
<td>0.555</td>
<td>0.551</td>
<td>0.595</td>
</tr>
</tbody>
</table>

**First Stage Results**

<table>
<thead>
<tr>
<th>Dependent Variable: 10-Yr. Difference in Log Share Ever Married</th>
<th>OLS</th>
<th>IV (1)</th>
<th>IV (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Log Cohort Size at Birth</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>R</em>^2</td>
<td>0.440***</td>
<td>0.765</td>
<td>0.805</td>
</tr>
</tbody>
</table>

**Significant at 5%, *** 1%. Robust std. 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). Includes all decennial cohorts born between 1910 and 1970, in all states except HI and AK. Sources: US Population Counts, 1910-1990; IPUMS USA, 1940-2010.

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

<table>
<thead>
<tr>
<th>Dependent Variable: N-Yr. Difference in Log Cohort Size</th>
<th>1-yr</th>
<th>3-yr</th>
<th>5-yr</th>
<th>7-yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ban ∗ Pill,c,s</td>
<td>0.012**</td>
<td>0.035***</td>
<td>0.041***</td>
<td>0.046**</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.011)</td>
<td>(0.015)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Ban,s</td>
<td>-0.002</td>
<td>-0.003</td>
<td>0.006</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.010)</td>
<td>(0.012)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>N</td>
<td>1248</td>
<td>1152</td>
<td>1056</td>
<td>960</td>
</tr>
</tbody>
</table>

* Significant at 10%, ** 5%, *** 1%. 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. Sources: 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)

<table>
<thead>
<tr>
<th>Dependent Variable: N-Yr. Difference in Log Share Ever Married (Men)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N-yr Difference</strong></td>
</tr>
<tr>
<td>in Log Cohort Size</td>
</tr>
<tr>
<td><em>R</em>^2</td>
</tr>
<tr>
<td>N</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Dependent Variable: N-Yr. Difference in Log Share Ever Married (Women)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N-yr Difference</strong></td>
</tr>
<tr>
<td>in Log Cohort Size</td>
</tr>
<tr>
<td><em>R</em>^2</td>
</tr>
<tr>
<td>N</td>
</tr>
</tbody>
</table>

* Significant at 10%, ** 5%, *** 1%. See note in Table 4.
### Table 6: Comstock Laws and Birth Weight

<table>
<thead>
<tr>
<th></th>
<th>Share with Birth Weight &lt; 1500</th>
<th>Share with Birth Weight &lt; 2500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ban * Pill&lt;sub&gt;c,s&lt;/sub&gt;</td>
<td>0.0001025 (.0002037)</td>
<td>-0.0004241 (.0007311)</td>
</tr>
<tr>
<td>Ban&lt;sub&gt;s&lt;/sub&gt;</td>
<td>-0.000462 (0.0003064)</td>
<td>0.0006242 (0.0006155)</td>
</tr>
<tr>
<td>N</td>
<td>1006</td>
<td>1006</td>
</tr>
</tbody>
</table>

See note in Table 4.

### Table 7: Calibrated Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment-Model:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>µ&lt;sub&gt;0&lt;/sub&gt;</td>
<td>Cost of investment</td>
<td>1.06</td>
</tr>
<tr>
<td>µ&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Cost of investment</td>
<td>0.07</td>
</tr>
<tr>
<td>µ&lt;sub&gt;2&lt;/sub&gt;</td>
<td>Cost of investment</td>
<td>3.90</td>
</tr>
<tr>
<td>θ&lt;sub&gt;h&lt;/sub&gt;</td>
<td>Prob. of meeting a woman</td>
<td>0.85</td>
</tr>
<tr>
<td>ρ&lt;sub&gt;0&lt;/sub&gt;</td>
<td>Probability of divorce</td>
<td>0.3</td>
</tr>
<tr>
<td>ρ&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Probability of divorce</td>
<td>0.47</td>
</tr>
</tbody>
</table>

δ-Model:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>α&lt;sub&gt;0&lt;/sub&gt;</td>
<td>Value of being single</td>
<td>-0.07</td>
</tr>
<tr>
<td>α&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Value of being single</td>
<td>0.50</td>
</tr>
<tr>
<td>ρ&lt;sub&gt;0&lt;/sub&gt;</td>
<td>Probability of divorce</td>
<td>-0.03</td>
</tr>
<tr>
<td>ρ&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Probability of divorce</td>
<td>1.03</td>
</tr>
</tbody>
</table>
Figure 1: 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 2: Share Never Married and Not Cohabiting By 30

(a) White

(b) Black

* See note in Figure 1.
Figure 3: Share Never Married and Not Cohabiting By 40

(a) White

(b) Black

* See note in Figure 1.
Figure 4: Growth by Region: States with Sales Bans - States without Sales Bans

A. Growth of Total Births

B. Growth of Total Adult Population, at Age 25


Figure 5: Share Never Married

A. Investment-Model

B. δ-Model
Figure 6: Share Ever Divorced

A. Data

B. Investment-Model

C. $\delta$-Model

Figure 7: Fraction of Young Men Who Invest

Figure 8: Age Difference Between Spouses and Cohort Size

Sources: Vital Statistics of the United States; IPUMS CPS, 1962-2011; IPUMS USA, 1950-1960. For details on how the variable age difference between spouses was constructed, see Data Appendix B.
A Appendix: Proofs and Derivations

A.1 Reservation Values

We begin by characterizing the decisions of a man of age 1 in period \( t \). If an old man chooses to be single in the second period, his lifetime utility takes the following form:

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

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

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

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 or until they divorce. Divorce can occur each period with a probability \( q = 1 - p \). If a couple divorces, each individual receives the value of being single for the remainder of their lifetime. 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:

\[
v_{0,0,t} = \eta \sum_{t=0}^{T} \beta^t p^t + 2 \delta \sum_{t=1}^{T} \beta^t (1 - p^t) = \eta \sum_{t=0}^{T} \beta^t p^t + 2 \delta \sum_{t=1}^{T} \beta^t - 2 \delta \sum_{t=1}^{T} \beta^t p^t =
\]

\[
\eta \sum_{t=0}^{T} \beta^t p^t + 2 \delta \sum_{t=1}^{T} \beta^t + 2 \delta - 2 \delta \sum_{t=1}^{T} \beta^t p^t - 2 \delta = \eta \sum_{t=0}^{T} \beta^t p^t + 2 \delta \sum_{t=0}^{T} \beta^t - 2 \delta \sum_{t=0}^{T} \beta^t p^t =
\]

\[
(\eta - 2 \delta) \sum_{t=0}^{T} \beta^t p^t + 2 \delta \sum_{t=0}^{T} \beta^t = \frac{1 - (p \beta)^{T+1}}{1 - p \beta} (\eta - 2 \delta) + \frac{1 - \beta^{T+1}}{1 - \beta} - 2 \delta,
\]

where the last equality follows from the following geometric series formula:

\[
\sum_{t=0}^{T} ab^t = a \frac{1 - b^{T+1}}{1 - b}.
\]

If the couple is composed of an older man and a woman, the man will die one period earlier. As a consequence, following the same steps as in the derivation of \( v_{0,0,t} \), their lifetime utility
takes the following form:

\[ v_{0,1,t} = \eta \sum_{t=0}^{T-1} \beta^t p^t + 2\delta \sum_{t=1}^{T-1} \beta^t (1 - p^t) + \beta^T = \frac{1 - (p\beta)^T}{1 - p\beta} \left( \eta - 2\delta \right) + \frac{1 - \beta^T}{1 - \beta} 2\delta + \beta^T \delta. \]

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,

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

where the parameter \( \gamma \in [0,1] \) allows for possible asymmetries in the way the marriage surplus is divided and \( v_{m,1,t} \) and \( v_{w,0,t} \) are the value of being single in this and future periods that were computed above. A similar equation can be derived for the woman.

We can solve the model starting with the decisions of a man of age 1 in period \( t \). With probability \( \theta_{1,t} \), he meets a woman and they marry if their joint lifetime utility from marrying \( v_{0,1,t} \) is greater than the sum of their lifetime utilities if they choose to stay single \( v_{1,t} + v_{0,t} \). As a consequence, they will marry if and only if

\[ \frac{1 - (p\beta)^T}{1 - p\beta} (\eta - 2\delta) + \frac{1 - \beta^T}{1 - \beta} 2\delta + \beta^T \delta - v_{1,t} - v_{0,t} \geq \frac{1 - \beta^T}{1 - \beta} 2\delta + \beta^T \delta. \]

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

\[ \eta_{1,t} = 2\delta. \]

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

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

As a consequence, the expected value function of an older man can be written in the following
form:
\[ v_{1,t}^m = \left( \delta \frac{1 - \beta^T}{1 - \beta} + E \left[ \gamma (\eta - 2\delta) \mid \eta \geq 2\delta \right] \frac{1 - (p\beta)^T}{1 - p\beta} \right) (1 - F (\eta_{1,t})) \theta_{1,t}^m + \]
\[ + \delta \frac{1 - \beta^T}{1 - \beta} F (\eta_{1,t}) \theta_{1,t}^m + \delta \frac{1 - \beta^T}{1 - \beta} (1 - \theta_{1,t}^m). \]

It is composed of three parts. The first term describes the value for the older 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 term characterizes the value of meeting a woman with a match quality \( \eta \) that is below the reservation value \( \eta_{1,t} \) times the probability of this event. Finally, the last term captures the value of not meeting a woman in the current period multiplied by the probability. By replacing \( \eta_{1,t} = 2\delta \) and simplifying some of the terms, we obtain the following equation for the value function:

\[ v_{1,t}^m = \delta \frac{1 - \beta^T}{1 - \beta} + E \left[ \gamma (\eta - 2\delta) \mid \eta \geq 2\delta \right] \frac{1 - (p\beta)^T}{1 - p\beta} (1 - F (2\delta)) \theta_{1,t}^m. \] (5)

We are now in position to consider the decision of a younger 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

\[ 2\delta \frac{1 - \beta^{T+1}}{1 - \beta} + (\eta - 2\delta) \frac{1 - (p\beta)^{T+1}}{1 - p\beta} \geq 2\delta + \beta v_{1,t+1}^m + \beta \delta \frac{1 - \beta^T}{1 - \beta}, \]

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. With this expression, we can now solve for the reservation value of a man of age 0. Substituting for the expected value function of an older man using equation (5) and simplifying some of the terms, we obtain the following equation for the reservation value of a younger man:

\[ \eta_{0,t} = 2\delta + \beta \frac{1 - (p\beta)^T}{1 - (p\beta)^{T+1}} \gamma \left\{ E [\eta \mid \eta \geq 2\delta] - 2\delta \right\} (1 - F (2\delta)) \theta_{1,t+1}^m. \] (6)

Using \( \eta_{0,t} \), one can derive the expected value function for a woman and a younger man. They are presented in Appendix A.6.

### A.2 Steady State

In this subsection, we use the reservation values discussed above to solve for the steady state equilibrium in the marriage market. To do this, we have to derive the probability that a younger man meets a woman \( \theta_{0,t}^m \) and the corresponding probability for an older man \( \theta_{1,t}^m \).
Let \( N_{a,t} \) be the number of individuals of gender \( i \), age \( a \), and period \( t \) who are present in the marriage market. Then \( \theta_{0,t}^m \) and \( \theta_{1,t}^m \) can be derived by noting that

\[
\theta_{0,t}^m = \theta_{1,t}^m = \frac{N_{0,t}^w}{N_{0,t}^m + N_{1,t}^m}.
\]

(7)

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

\[
N_{m,1,t} = N_{0,t-1}^m (1 - \theta_{0,t-1}^m) + N_{0,t-1}^m \theta_{0,t-1}^m F(\eta_{0,t-1}) = N_{0,t-1}^m (1 - \theta_{0,t-1}^m (1 - F(\eta_{0,t-1}))).
\]

(8)

We can now replace for \( \theta_{0,t-1}^m \) using (7) and obtain the following equation for \( N_{m,1,t}^m \):

\[
N_{m,1,t} = N_{0,t-1}^m \left( 1 - \frac{N_{0,t-1}^w}{N_{0,t-1}^m + N_{1,t-1}^m} \left( 1 - F(\eta_{0,t-1}) \right) \right) = N_{0,t-1}^m \left( \frac{N_{0,t-1}^m + N_{1,t-1}^m - N_{0,t-1}^w \left( 1 - F(\eta_{0,t-1}) \right)}{N_{0,t-1}^m + N_{1,t-1}^m} \right).
\]

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

\[
N_{1,t}^m = N_0^m \left( \frac{N_0^m + N_1^m - N_0^w (1 - F(\eta_0))}{N_0^m + N_1^m} \right).
\]

We can now solve for \( N_{1,t}^m \) and obtain

\[
N_{1,t}^m = \sqrt{(N_0^m)^2 - N_0^m N_0^w + N_0^m N_0^w F(\eta_0)}.
\]

Generally, men and women have identical cohort size, i.e. \( N_{0,t}^m = N_{0,t}^w = N_{0,t} \). In this case the solution for \( N_{1,t}^m \) simplifies to

\[
N_{1,t}^m = N_0 F(\eta_0)^{\frac{1}{2}}.
\]

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

$$\theta^m_0 = \theta^m_1 = \frac{N^w_0}{N^m_0 + \sqrt{(N^m_0)^2 - N^m_0 N^w_0 + N^m_0 N^w_0 F(\eta_0)}}.$$  

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

$$\theta^m_0 = \theta^m_1 = \frac{N_0}{N_0 + N_0 F(\eta_0)} = \frac{1}{1 + F(\eta_0)}.$$  

To determine the reservation value of younger men in steady state, we can substitute for $\theta^m_1$ in the equation that determines the reservation value (6). 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:

$$\eta_{ss} = 2\delta + \beta \left( \frac{1 - (p\beta)^T_1}{1 - (p\beta)^T_{\tau+1}} \right) \gamma \{ E[\eta | \eta \geq 2\delta] - 2\delta \} \left( 1 - F(2\delta) \right) \frac{N^w_0}{N^m_0 + \sqrt{(N^m_0)^2 - N^m_0 N^w_0 + N^m_0 N^w_0 F(\eta_{ss})}},$$

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

$$\eta_{ss} = 2\delta + \beta \left( \frac{1 - (p\beta)^T_1}{1 - (p\beta)^T_{\tau+1}} \right) \gamma \{ E[\eta | \eta \geq 2\delta] - 2\delta \} \left( 1 - F(2\delta) \right) \frac{1}{1 + F(\eta_{ss})^{1/2}}.$$  

Note that $F(\eta)$ is monotonically increasing in $\eta$. As a consequence, there is a unique solution for $\eta_{ss}$. Moreover, if men and women have identical cohort sizes, the steady state reservation value is independent of $N^m_0$ and $N^w_0$.

A.3 Proof of Proposition 2 and The Effect of an Unexpected Shock to Cohort Size

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 $N_0 + \Delta$. According to equation (7), the probabilities $\theta^m_j, i$ take the following form:

$$\theta^m_{0,t} = \theta^m_{1,t} = \frac{N_{0,t}}{N_{0,t} + N^m_{1,t}} \quad \text{if } t < \tau$$

and

$$\theta^m_{0,t} = \theta^m_{1,t} = \frac{N_{0,t} + \Delta}{N_{0,t} + \Delta + N^m_{1,t}} \quad \text{if } t \geq \tau.$$  

Consider the period in which the shock is realized and notice that $N^m_{1,\tau}$ are the men born in period $\tau - 1$ who did not marry when younger. As a consequence, $N^m_{1,\tau}$ equals the number of older men in steady state, i.e. $N^m_{1,\tau} = N_{0,\tau-1} F(\eta_{ss})^{1/2} = N_0 F(\eta_{ss})^{1/2}$. Substituting for $N^m_{1,\tau}$
in the probabilities \( \theta_{m,t}^{m} \), we have that in period \( \tau \)

\[
\theta_{0}^{m} = \theta_{1}^{m} = \frac{N_{0} + \Delta}{N_{0} + \Delta + N_{0}F(\eta_{ss})} = \frac{1}{1 + \frac{N_{0}}{N_{0} + \Delta} F(\eta_{ss})^2}.
\]

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 younger man, \( \theta_{w}^{m} = \frac{N_{0,t}}{N_{0,t} + N_{1,t}} \), is equivalent to the probability that a man meets a woman. Therefore, the previous result also implies that a positive cohort shock increases the probability that a woman meets a younger man.

We can now determine the effect of a shock to cohort size on the reservation value of younger men \( \eta_{0,\tau} \). Notice that in the determination of \( \eta_{0,\tau} \) a younger 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 older men at \( \tau + 1 \), which can be written as follows:

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

Using equation (8), we can substitute for \( N_{1,\tau + 1} \) to obtain the following expression:

\[
\theta_{0,\tau + 1} = \theta_{1,\tau + 1} = \frac{N_{0} + \Delta}{N_{0} + \Delta + N_{1,\tau + 1}} = \frac{1}{1 + \left(1 - \theta_{0,\tau} (1 - F(\eta_{0,\tau}))\right)}.
\]

We can now substitute for \( \theta_{1,\tau + 1} \) in the equation that determines \( \eta_{0,\tau} \) to obtain

\[
\eta_{0,\tau} = 2 \delta + \beta \frac{1 - (p \beta)^{T}}{1 - (p \beta)^{T + 1}} \gamma \left\{ E[\eta | \eta \geq 2 \delta] - 2 \delta \right\} \frac{1}{1 + \left(1 - \theta_{0,\tau} (1 - F(\eta_{0,\tau}))\right)}.
\]

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

\[
\eta_{0,ss} = 2 \delta + \beta \frac{1 - (p \beta)^{T}}{1 - (p \beta)^{T + 1}} \gamma \left\{ E[\eta | \eta \geq 2 \delta] - 2 \delta \right\} \frac{1}{1 + \left(1 - \theta_{0,ss} (1 - F(\eta_{0,ss}))\right)}.
\]

Earlier in this section we have shown that, with a positive shock to cohort size, \( \theta_{0,\tau} > \theta_{0,ss} \).
As a consequence, a simple comparison of the last two equations implies that an increase in cohort size has the effect of increasing the reservation value of younger men. Specifically, by substituting \( \theta_{0,ss}^{m} \) with \( \theta_{0,\tau}^{m} \) and by using the result that \( \theta_{0,\tau}^{m} > \theta_{0,ss}^{m} \), we obtain the following
inequality:

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

Since the left hand side of the inequality is increasing in \( \eta_{0,ss} \) and the right hand side is decreasing in \( \eta_{0,ss} \), equation (9) implies that \( \eta_{0,ss} > \eta_{0,ss} \).

A.4 Proof of Proposition 3

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

\[ P(\text{woman marries at } \tau) = \theta^w_{0,\tau} (1 - F(\eta_{0,\tau})) + (1 - \theta^w_{0,\tau}) (1 - F(2\delta)) \]

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

\[ P(\text{woman marries at } \tau) = \theta^w_{0,\tau} (1 - F(\eta_{0,\tau})) + (1 - \theta^w_{0,\tau}) (1 - F(2\delta)) \]

\[ = \theta^w_{0,ss} (1 + \phi_\tau) + (1 - \theta^w_{0,ss}) (1 + \lambda_\tau) + (1 - \theta^w_{0,ss} (1 + \phi_\tau)) (1 - F(2\delta)) \]

\[ = \theta^w_{0,ss} (1 - F(\eta_{0,ss})) + (1 - \theta^w_{0,ss}) (1 - F(2\delta)) - \theta^w_{0,ss} \lambda_\tau F(\eta_{0,ss}) + \theta^w_{0,ss} \phi_\tau (1 - F(\eta_{0,ss}) (1 + \lambda_\tau)) \]

\[ - \theta^w_{0,ss} \phi_\tau (1 - F(2\delta)) \]

\[ = P(\text{woman marries at } ss) - \theta^w_{0,ss} \lambda_\tau F(\eta_{0,ss}) + \theta^w_{0,ss} \phi_\tau (1 - F(\eta_{0,ss})) - \theta^w_{0,ss} \phi_\tau (1 - F(2\delta)) \]

\[ < P(\text{woman marries at } ss) - \theta^w_{0,ss} \lambda_\tau F(\eta_{0,ss}) \]

\[ < P(\text{woman marries at } ss). \]

A.5 Proof of Proposition 4

We prove the Proposition in two steps. We first prove that the probability that a man marries when younger in period \( t \) increases with cohort size. When then prove that a man marries when younger or older increases with cohort size.

First step. Let \( P^m_t \) be the probability that a man marries when younger the period of the shock \( \tau \). Since we consider the case of a permanent shock to cohort size we have
\[ N^m_{0,\tau} = N^m_{0,\tau+1}. \] Using equation (8), we can therefore write the number of older men in period \( t + 1 \) as follows:

\[ N^m_{0,\tau+1} = N^m_{0,\tau} \left( 1 - \theta^m_{0,\tau} \left( 1 - F \left( \eta_0,\tau \right) \right) \right) = N^m_{0,\tau} \left( 1 - P^{ym}_\tau \right). \tag{11} \]

Using the previous equation and equation (9), the reservation utility of a younger man can be written in the following form:

\[ \eta_{0,\tau} = A + B N^m_{0,\tau} N^m_{0,\tau} + N^m_{0,\tau+1} \left( 1 - \theta^m_{0,\tau} \left( 1 - F \left( \eta_0,\tau \right) \right) \right) = A + B \frac{1}{2 - P^{ym}_\tau} = A + B \frac{1}{2 - P^{ym}_\tau}. \]

Proposition 2 establishes that \( \eta_{0,ss} < \eta_{0,\tau} \). Hence,

\[ \eta_{0,ss} = A + B \frac{1}{2 - P^{ym}_{ss}} < A + B \frac{1}{2 - P^{ym}_\tau} = \eta_{0,\tau}. \]

The inequality implies that \( P^{ym}_{ss} > P^{ym}_\tau \). We can therefore conclude that an increase in cohort size increases the probability that a man marries when younger.

Second step. The probability that a man marries when younger or older \( P^m_\tau \) can be written as the probability that a man married when younger in period \( \tau \) plus the probability that the same man marries when older in period \( \tau + 1 \), i.e.

\[ P^m_\tau = \theta^m_{0,\tau} \left( 1 - F \left( \eta_0,\tau \right) \right) + \left( 1 - \theta^m_{0,\tau} \left( 1 - F \left( \eta_0,\tau \right) \right) \right) \theta^m_{1,\tau+1} \left( 1 - F \left( 2\delta \right) \right). \tag{12} \]

The first part of the right hand side is the probability that a younger man meets a woman and marries her in period \( \tau \), which we denoted with \( P^{ym}_\tau \). The second part is the probability that a younger man does not marry in period \( \tau \), \( 1 - P^{ym}_\tau \), meets a woman in period \( \tau + 1 \), and marries her. Using equation (11), the probability that an older men meets a woman can be written as follows:

\[ \theta^m_{1,\tau+1} = \frac{N^m_{0,\tau+1}}{N^m_{0,\tau+1} + N^m_{1,\tau+1}} = \frac{1}{2 - P^{ym}_\tau}. \]

As a consequence, equation (12) can be written as follows:

\[ P^m_\tau = P^{ym}_\tau + \frac{1 - P^{ym}_\tau}{2 - P^{ym}_\tau} \left( 1 - F \left( 2\delta \right) \right). \]

Taking the derivative with respect to cohort size \( N \) of both size and rearranging terms, we have,

\[ \frac{\partial P^m_\tau}{\partial N} = \frac{\partial P^{ym}_\tau}{\partial N} \left[ 1 - \frac{1 - F \left( 2\delta \right)}{\left( 2 - P^{ym}_\tau \right)^2} \right] > \frac{\partial P^{ym}_\tau}{\partial N} F \left( 2\delta \right) > 0, \]

where the first inequality follows from \( \left( 2 - P^{ym}_\tau \right)^2 > 1 \) and the second from the first step of
the proof. Hence, an increase in cohort size increases the probability that a man marries.

A.6 Expected Value Functions

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

\[
v_{0,t}^m = \theta_{0,t}^m \left(1 - F(\eta_{0,t})\right) \left\{\delta + \beta v_{1,t}^m + \gamma \left\{2\delta \frac{1 - \beta^{T+1}}{1 - \beta} + \frac{1 - (p\beta)^{T+1}}{1 - (p\beta)} E[\eta - 2\delta | \eta \geq \eta_{0,t}] - \delta - \beta v_{1,t}^m \right\}\right\} + \theta_{0,t}^m F(\eta_{0,t}) \left(\delta + \beta v_{1,t}^m\right) + \left(1 - \theta_{0,t}^m\right) \left(\delta + \beta v_{1,t}^m\right).
\]

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 woman 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 younger times the probability.

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

\[
v_{0,t}^w = \theta_{0,t}^w \left(1 - F(\eta_{0,t})\right) \left\{\frac{1 - \beta^{T+1}}{1 - \beta}\delta + (1 - \gamma) \left\{2\beta \frac{1 - \beta^{T+1}}{1 - \beta} + \frac{1 - (p\beta)^{T+1}}{1 - (p\beta)} E[\eta - 2\delta | \eta \geq \eta_{0,t}] - \delta - \beta^T \delta\right\}\right\} + \theta_{0,t}^w F(\eta_{0,t}) \frac{1 - \beta^{T+1}}{1 - \beta} + \left(\delta + \beta v_{1,t}^w\right)\right\} + \theta_{0,t}^w F(\eta_{0,t}) \frac{1 - \beta^{T+1}}{1 - \beta} \delta + \left(1 - \theta_{0,t}^w\right) \left(\delta + \beta v_{1,t}^w\right) + \left(1 - \theta_{0,t}^w\right) \left(\delta + \beta v_{1,t}^w\right) + \beta v_{1,t}^w.
\]

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

A.7 Basic Search Model with Divorce

In Figure A1, we show simulation results for the basic search model, under two different assumptions about the probability with which a divorce may occur. Under the first assumption, the probability that a divorce occurs is constant. Under the second assumption, the probability of divorce is modeled as a decreasing function of match quality of the form \(q(\eta) = \gamma_0 + \gamma_1 \eta\). As Figure A1 shows, under both assumptions, the share never married changes in opposite directions for men and women, in contrast to the patterns observed in
the data. Additionally, as discussed in the paper, the share never married is higher for all cohorts when the probability of divorce is decreasing in the value of marriage since younger men become more selective and increase their reservation utility.

**Figure A1**: Share of Men and Women Never Married, Basic Search Model
B Appendix: Data Description

Table B1 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 B1, when cohort size is computed for the U.S. population there is little difference between cohort size at birth and cohort size 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 B1, 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 of individuals between the ages of 0 and 4 and 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 of 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.

Table B1: Data Sets Used in the Construction of the Main Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variation of Interest</th>
<th>Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>State Level, Decennial Years</td>
<td>U.S. Decennial Census, 1910-1970</td>
</tr>
<tr>
<td></td>
<td>State-Level, Decennial Years</td>
<td>U.S. Decennial Census, 1930-1990</td>
</tr>
<tr>
<td>Age Differences of Spouses</td>
<td>National, Yearly</td>
<td>U.S. Decennial Census, 1940-2000</td>
</tr>
<tr>
<td></td>
<td>State Level, Decennial Years</td>
<td>U.S. Decennial Census, 1970-2000</td>
</tr>
</tbody>
</table>

Figure B1: Cohort Size at Birth and at Age 30