The Baby Boom and World War II: A Macroeconomic Analysis*

Matthias Doepke
Northwestern University

Moshe Hazan
Hebrew University

Yishay D. Maoz
Open University of Israel

June 2012

Abstract

We argue that one major cause of the U.S. postwar baby boom was the rise in female labor supply during World War II. We develop a quantitative dynamic general equilibrium model with endogenous fertility and female labor-force participation decisions. We use the model to assess the impact of the war on female labor supply and fertility in the decades following the war. For the war generation of women, the high demand for female labor brought about by mobilization leads to an increase in labor supply that persists after the war. As a result, younger women who turn adult in the 1950s face increased labor-market competition, which impels them to exit the labor market and start having children earlier. The effect is amplified by the rise in taxes necessary to pay down wartime government debt. In our calibrated model, the war generates a substantial baby boom followed by a baby bust.

*We thank the editor, three anonymous referees, Stefania Albanesi, Leah Boustan, Larry Christiano, Alon Eizenberg, Eric Gould, Jeremy Greenwood, Christian Hellwig, Lee Ohanian, as well as participants at many conference and seminar presentations for helpful comments. David Lagakos, Marit Hinnosaar, Amnon Schreiber, and Veronika Selezneva provided excellent research assistance. Financial support by the Maurice Falk Institute for Economic Research in Israel, the National Science Foundation (grant SES-0217051), and the Alfred P. Sloan is gratefully acknowledged. Doepke: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (e-mail: doepke@northwestern.edu). Hazan: Department of Economics, Hebrew University of Jerusalem, Mt. Scopus, Jerusalem 91905 (e-mail: moshe.hazan@huji.ac.il). Maoz: Department of Management and Economics, The Open University of Israel, 1 University Road, Raanana 43107 (e-mail: yishayma@openu.ac.il).

All the day long, whether rain or shine,
She's a part of the assembly line.
She's making history, working for victory,
Rosie the Riveter.[1]

1 Introduction

In the two decades following World War II the United States experienced a massive baby boom. The total fertility rate[2] increased from 2.3 in 1940 to a maximum of 3.8 in 1957 (see Figure 1). Similarly, the data on cohort fertility show an increase from a completed fertility rate[3] of about 2.4 for women whose main childbearing period just preceded the baby boom (birth cohorts 1911–1915) to a rate of 3.2 for the women who had their children during the peak of the baby boom (birth cohorts 1930–1935; see Figure 2). The baby boom was followed by an equally rapid baby bust. The total fertility rate fell sharply throughout the 1960s, and was below 2.0 by 1973. The baby boom constituted a dramatic, if temporary, reversal of a century-long trend towards lower fertility rates. Understanding the causes of the baby boom is thus is a key challenge for demographic economics.

In this paper, we propose a novel explanation of the baby boom, based on the demand for female labor during World War II. As documented by Acemoglu, Autor, and Lyle (2004), the war induced a large positive shock to the demand for female labor. While men were fighting the war in Europe and Asia, millions of women were drawn into the labor force and replaced men in factories and offices.[4] The effect of the war on female employment was not only large, but also persistent: the women who worked during the war accumulated valuable labor-market experience, and consequently many of them continued to work after the war.

[1] "Rosie the Riveter," lyrics by Redd Evans and John Jacob Loeb, 1942.

[2] The total fertility rate in a given year is the sum of age-specific fertility rates over all ages. It can be interpreted as the total number of children an average woman will have over her lifetime if age-specific fertility rates stay constant over time.

[3] The completed fertility rate is the average lifetime number of children born to mothers of a specific cohort. Dynamic patterns of total and completed fertility rates can deviate if there are shifts in the timing of births across cohorts.

[4] The U.S. government actively campaigned for women to join the war effort. "Rosie the Riveter," a central character in the wartime campaign for female employment, has become a cultural icon and a symbol of women's expanding economic role.

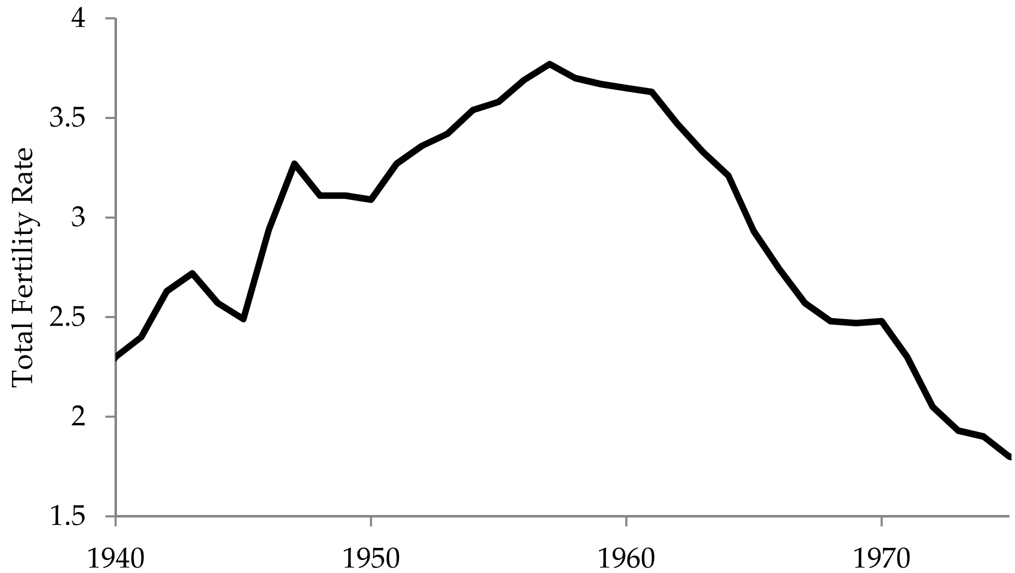

Figure 1: The Total Fertility Rate in the United States (Source: Chesnais 1992)

At first sight, it might seem that this additional supply of female labor should generate the opposite of a baby boom: women who work have less time to raise children and usually decide to have fewer of them. The key to our argument, however, is that the one-time demand shock for female labor had an *asymmetric* effect on different cohorts of women. The only women who stood to gain from additional labor market experience were those who were old enough to work during the war. For younger women who were still in school during the war the effect was negative: when they turned adult after the war and entered the labor market, they faced increased competition. In addition to the men who returned from the war, a large number of the experienced women of the war generation were still in the labor force. We argue that this led to less demand for inexperienced young women, who were crowded out of the labor market and chose to have more children instead. It is these younger women who account for the bulk of the baby boom.

Our explanation is consistent with the observed patterns of female labor-force participation before the war and during the baby boom. In the years leading up to the war, the vast majority of single women in their early 20s were working. In contrast, labor-force participation rates for married women were low. Hence,

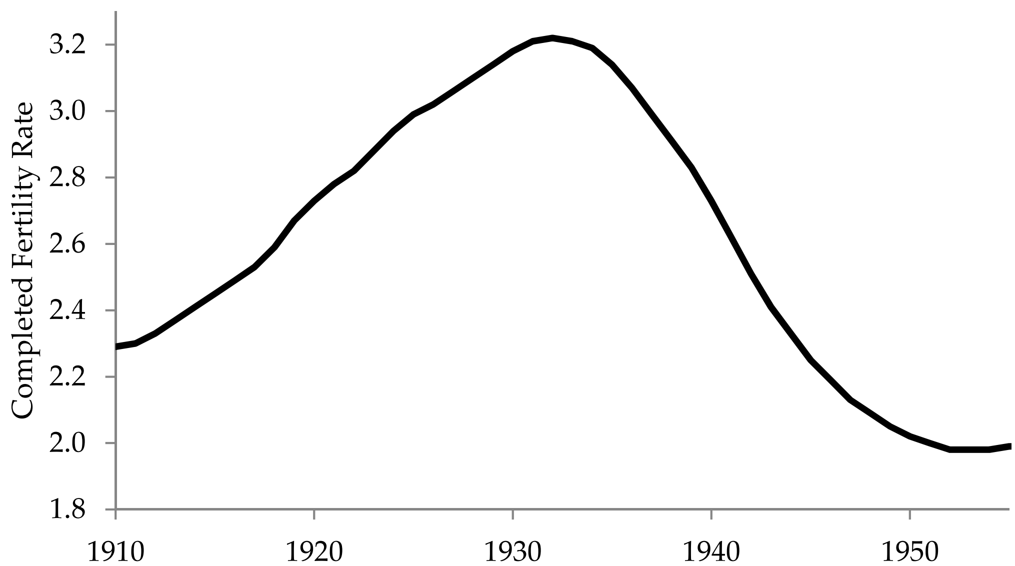

Figure 2: The Completed Fertility Rate in the United States by Cohort (i.e., by Birth Year of Mother. Source: Observatoire Démographique Européen)

a typical woman would enter the labor force after leaving school, and then quit working (usually permanently) once she got married and started to have children. Figure 3 shows how the labor supply of young (ages 20–32) and older (ages 33–60) women evolved after the war. During the baby-boom period, the labor supply of older women increased sharply, whereas young women worked less. A substantial part of the drop of young female labor supply is due to a compositional shift from single to married women. On average, these women decided to get married younger than earlier cohorts had done, which (given the low average labor supply of married women) lowered the total amount of labor supplied by young women. Our theory generates the same pattern as a result of the wartime demand shock for female labor.

We interpret the decline in young women's labor supply as a crowding-out effect due to higher participation by older women. This interpretation is consistent with the observation that the relative wages of young women declined during the baby boom period. Figure 4 displays the wages of single women aged 20–24 relative to the wages of men in the same age group. Relative female wages decline in both 1950 and 1960, and recover strongly only in 1970 during the baby

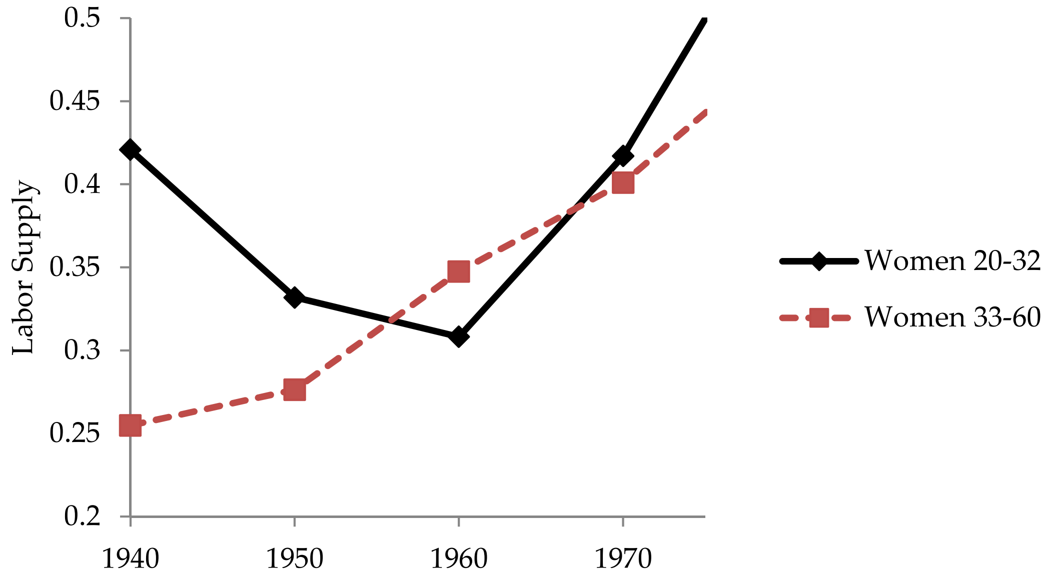

Figure 3: Labor Supply by Young (20–32) and Old (33–60) Women in the United States relative to Men in the same Age Group (includes women of all races and marital statuses; see Appendix A.3 for details)

bust.[5] Our theory reproduces these relative-wage shifts through the war-induced increase in the labor-force participation of older women. In contrast, a model in which young women withdraw from the labor market for other reasons would predict that relative female wages should have risen in the baby-boom period.

Our theory is also supported by the observation that most of the baby boom is accounted for by young mothers. Data on age-specific fertility[6] show that both in absolute and relative terms, women aged 20–24 experienced the largest increase in fertility, with almost a doubling of fertility between 1940 and the peak of the baby boom. For women aged 25–29 the increase in fertility is more than one-third smaller than in the younger group, and among even older mothers the baby boom is either small or nonexistent. In line with these numbers, the average age at first birth dropped by more than 1.5 years from 1940 to the late 1950s.

[5]Notice that the decline in relative wages for *young* women does not imply that the *overall* gender gap widens. In fact, average relative wages for all working women rise from 1940 to 1950 and 1960. However, this is due to a compositional shift from young women to a labor force consisting of a larger fraction of older women with work experience. In our theory, it is the relative wage of young women that drives fertility decisions.

[6]Number of Births per 1,000 Women in Different Age Groups in the United States, from Vital Statistics of the United States, 1999, Volume I, Natality (Table 1–7).

Figure 4: Ratio of Average Female to Average Male Wages for Singles aged 20–24, 1940=100 (Source: U.S. Census; see Appendix A.3 for details)

In our theory, fertility increases because women exit the labor force and start having children earlier, which implies that, as observed in the data, the increase in fertility takes place at the beginning of the child-bearing period.

To provide empirical evidence for the proposed mechanism, we follow the approach of Acemoglu, Autor, and Lyle (2004) of using variation in mobilization rates across states to identify the effect of the war. In line with the first part of our hypothesis, Acemoglu, Autor, and Lyle show that the wartime increase in female labor supply led to a persistent increase in the labor force participation of older women as well as lower relative female wages. Building on these results, we show that states with a greater mobilization of men during the war (and thus a higher wartime demand for female labor) also had a larger postwar increase in fertility. In addition, in high-mobilization states young women were less likely to work and more likely to be married during the baby-boom period. These are exactly the relationships predicted by our hypothesis.

We then develop a dynamic general-equilibrium model to demonstrate that the labor-market mechanism outlined above can account for much of the increase in fertility during the baby boom. In addition, the model allows us to consider additional driving forces of the baby boom that do not vary across states (in particular changes in taxation) and to evaluate whether our theory can explain the timing of the baby boom and baby bust. The model focuses on married couples' life-cycle decisions on fertility and female labor-force participation. In the model, all women start out working when young, but ultimately quit the labor force in order to have children. Since the fecund period is limited, having more children requires leaving the labor market earlier. Due to the time cost of having children and an adjustment cost of reentering the labor market, only some women resume work after having children. Since fertility and labor-force participation decisions are discrete, the model incorporates preference heterogeneity to generate heterogeneous behavior in these dimensions. At the aggregate level, the model features a standard production technology with limited substitutability of male and female labor. We calibrate the model to U.S. data, and then shock the model's balanced growth path with World War II, represented as a shock to government spending, a reduction in male labor supply, and an increase in female labor supply.

We find that the model does an excellent job at reproducing the main qualitative features of the U.S. baby boom: the patterns for fertility, the timing of births, female labor-force participation rates, and relative female wages are all consistent with empirical observations. The model does particularly well at reproducing the timing of the baby boom and baby bust. The baby boom reverses once the war generation of working women starts to retire from the labor market. This model implication results in a sharp reduction in fertility 15 to 20 years after the war shock, which closely matches the baby bust period of the 1960s.

Turning to quantitative implications, we find that in our baseline calibration the model can account for a major fraction of the increase in cohort fertility during the baby boom. The model generates a maximum increase in fertility of 0.6 children per woman, which compares to a maximum of 0.8 in the data. About 80 percent of the increase in fertility generated by the model is due to a crowding-out effect generated by higher labor-force participation of the war generation of women, with the remainder accounted for by the fiscal consequences of the war. The model also closely tracks the actual changes in labor supply by younger women throughout the baby boom period, and is consistent with the magnitude of changes in relative female wages.

Another way to assess the empirical relevance of the labor-market mechanism is to consider data on the baby boom in countries other than the United States. Most industrialized countries experienced a baby boom after World War II, but only some of them also underwent a substantial mobilization of female labor during the war. Our theory predicts that countries with a larger wartime increase in the female labor force should also experience larger baby booms. The international data is consistent with this prediction. In particular, we compare the baby boom in countries that had a wartime experience similar to the United States (Allied countries that mobilized for the war but did not fight on their own soil, namely Australia, Canada, and New Zealand) with neutral countries that did not experience a large demand shock for female labor (Ireland, Portugal, Spain, Sweden, and Switzerland). We find that the Allied countries experienced a large baby boom quite similar to the one in the United States, whereas the increase in fertility was much smaller in the neutral countries. We regard the larger baby boom in the Allied countries as a strong indication that our mechanism is relevant. At the same time, the fact that the neutral countries had baby booms at all also suggests that our mechanism cannot be the only explanation: some factor other than the dynamics of the female labor market must have also played a role.

The remainder of the paper is organized as follows. In the following section, we provide empirical evidence on the effect of wartime mobilization on fertility during the U.S. baby boom. The model economy is described in Section 3. Our main findings are presented in Section 4, where we discuss the model's quantitative implications for the effect of World War II on post-war fertility. International evidence is discussed in Section 5. In Section 6 we relate our results to the existing literature on the baby boom, and Section 7 concludes.

2 Evidence from Mobilization Rates

In a seminal contribution, Acemoglu, Autor, and Lyle (2004) use variation in mobilization rates across U.S. states to document the impact of the war on the labor market for women. The authors show that U.S. states with a greater mobilization of men during the war (and thus a higher demand for female labor) also had a larger postwar increase in female employment, whereas relative female wages declined relative to states with lower mobilization rates. These results confirm the link between the rise of female employment in World War II and increased subsequent competition in the female labor market that is an essential ingredient of our mechanism.

In this section, we build directly on Acemoglu, Autor, and Lyle and establish that states that had high mobilization rates during the war subsequently also experienced higher fertility and lower labor force participation by young women (those turning adult after the war). Figure 5 displays a cross plot of state mobilization rates for World War II and the change in fertility from 1940 to 1960. The measure of fertility (computed from census data) is the average number of own children under age 5 living in the household for women of ages 25–35. The fertility measure corresponds to births that occurred between 1935 and 1940 for the 1940 census and between 1955 and 1960 for the 1960 census, which covers the peak of the baby boom. The figure reveals a clear positive association between mobilization and the change in fertility. A regression of the fertility change on mobilization rate gives a coefficient of 0.723 with a t-statistic of 2.11. The size of the coefficient is economically and demographically significant: multiplying the regression coefficient with the average mobilization rate of 47.8 percent results in a fertility change of 0.35, which would account for most of the increase in this measure throughout the baby boom (from 1940 to 1960 the fertility measure increased by 0.37, see Table 1).

Of course, correlation does not imply causation, and it is possible that states differed in other dimensions that are correlated with mobilization rates and that also affected fertility. To deal with such concerns, we now examine the link between wartime mobilization and fertility in more detail.

2.1 Data Sources

For data on fertility, labor supply, and other individual characteristics we use the 1 percent Integrated Public Use Microdata Series (IPUMS) from the 1940 and 1960 censuses (Ruggles et al. 2010). We use data from the 48 contiguous states (Alaska

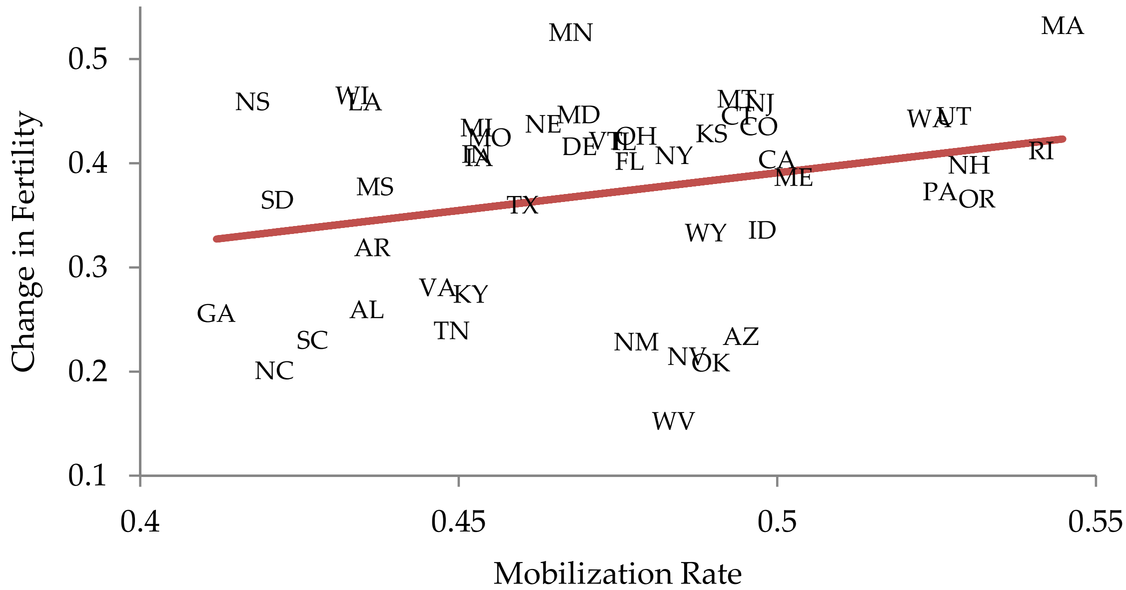

Figure 5: State Mobilization Rates for World War II and Change in Fertility from 1940 to 1960 (Average Number of Own Children Under Age 5 in Household for Women of Ages 25–35)

and Hawaii did not gain statehood until the 1950s) and also omit Washington, D.C. We exclude women living in group quarters. As the main fertility measure for a woman we use the number of own children under the age of 5 living in the same household. For labor supply, we consider a dummy variable representing whether a woman is currently employed, as well as the number of weeks worked in a year. We also consider information on marital status, namely an indicator of ever having been married (i.e., currently married, widowed, divorced, or separated), because in the data the beginning of childbearing is closely associated with marriage. We distinguish two different age groups, namely women aged 25–35 as the "young" group and women aged 45-55 as the "old group." In line with our theoretical ideas, the young age range is chosen such that in 1960 women in this group are at the peak of their fertility.[7] In addition, the women in this age group were between 10 and 20 years old at the end of World War II, so

[7]Notice that the fertility measure picks up birth in the preceding 5 years, thus starting at age 20 for the youngest women. The baby boom had only a small effect on fertility rates before age 20.

Table 1: Mean and Standard Deviation (in Parentheses) of Fertility, Marriage, and Labor Supply in the 1940 and 1960 Censuses

Variable	1940		1960	
	Age 25–35	Age 45–55	Age 25–35	Age 45–55
Children Under Age 5	0.49		0.86	
	(0.75)		(0.96)	
Ever Married	0.83		0.92	
	(0.38)		(0.21)	
Employed	0.28	0.18	0.32	0.43
	(0.45)	(0.38)	(0.47)	(0.50)
Weeks Worked/Year	15.4	10.4	15.4	20.6
	(22.5)	(20.0)	(20.9)	(22.8)

that they were mostly too young for the war to directly affect their labor supply.[8] In contrast, the older group sampled in the 1960 census was between 30 and 40 years old, an age range in which the war had a large direct effect on labor force participation.

Table 1 displays the mean and standard deviations for our main variables of interest. Fertility increased strongly from 1940 to 1960, with the mean number of own children under age 5 in a woman's household increasing from 0.49 to 0.86. Young women were also more likely to be married in 1960. For labor supply, there is little change for young women of ages 25–35 between 1940 and 1960, but large increase in the labor supply of older women, with more than a doubling in employment.

For mobilization rates we use the same variable as Acemoglu, Autor, and Lyle (2004), which is the fraction of registered men between the ages of 18 and 44 who were drafted or enlisted for war, by state.[9] The mobilization rates vary between 41.2 and 54.5 percent, with an average of 47.8 percent.

[8]While some of these women would have worked at the end of the war in their late teens, women in this age group were likely to work even before the war. Our results are robust to further reducing the age group to only include women who were minors at the end of the war.

[9]We thank the authors for making the data available to us.

2.2 Results

Our main results are based on individual-level regressions of the form:[10]

$$y_{ist} = \lambda_s + \pi d_{1960} + X'_{ist}\omega + \mu d_{1960} m_s + \epsilon_{ist}$$

using pooled census data from 1940 and 1960. Here y_{ist} is an outcome variable of interest (fertility, labor supply, or marriage), λ_s is a state fixed effect, d_{1960} is a dummy for 1960, X_{ist} is a vector of individual-level controls, and m_s is the state mobilization rate for World War II. The main parameter of interest is μ, the interaction of mobilization with the 1960 dummy. For example, in a fertility regression a positive estimate for μ would indicate that fertility went up by more between 1940 and 1960 in states with high mobilization rates than in states with low mobilization rates.

Table 2 displays results for the fertility, labor supply, and marriage decisions of young women. Each entry in Table 2 shows the estimate of the interaction term μ for a different specification. All regressions include dummy variables for observation year, age, race, state of residence, and state/country of birth. In the fertility regressions, we also control for the number of children older than 5. All the indicator variables except state/country of birth and state of residence are also interacted with the 1960 dummy in order to allow the effects to differ across the two periods.

Column 1 in Table 2 displays results for our most parsimonious specification. We find that women in states with high mobilization rates had substantially more children, worked less, and were more likely to be married than women in states with low mobilization rates. The parameter estimates are all highly statistically significant and imply a large quantitative impact of mobilization. For example, multiplying the coefficient on fertility with the average mobilization rate of 47.8 percent implies that wartime mobilization increased fertility (in terms of children under 5 years of age) by more than one-half of a child per woman in 1960. Similarly, evaluated at the average mobilization rate labor supply declines by more

[10]The empirical setup broadly follows Acemoglu, Autor, and Lyle (2004), see their regression equation (8). However, we focus on different outcome variables and there are also some differences in controls.

Table 2: Impact of WWII Mobilization Rates on Fertility, Labor Supply, and Marriage of Women Aged 25–35 (Coefficient Estimates from OLS and 2SLS Regressions for Variable "Mobilization Rate × 1960")

Dependent Variable	Regression				
	(1)	(2)	(3)	(4)	(5)
	Age 25–35 (N = 243554)				
Children under Age 5	1.146	0.665	0.573	2.277	2.050
	(0.259)	(0.232)	(0.208)	(0.381)	(0.347)
R^2	0.115	0.119	0.196	0.119	0.195
Employed	-0.820	-0.395	-0.204	-0.567	-0.229
	(0.149)	(0.129)	(0.096)	(0.192)	(0.138)
R^2	0.020	0.046	0.204	0.046	0.204
Weeks Worked	-27.756	-14.491	-5.429	-19.809	-3.666
	(8.164)	(7.591)	(5.578)	(10.782)	(7.908)
R^2	0.015	0.031	0.158	0.031	0.158
Ever Married	0.384	0.377		0.651	
	(0.119)	(0.122)		(0.179)	
R^2	0.046	0.063		0.063	
p-value First Stage				<0.0001	<0.0001
Education and Farm Controls	no	yes	yes	yes	yes
Marital Status Controls	no	no	yes	no	yes

Notes: Standard errors (in parentheses) are adjusted for clusters of state of residence and year of observation. Estimates are from separate regressions of pooled micro data from 1940 and 1960 census. Regressions 1-3 are OLS and 4-5 are 2SLS. Each outcome variable is regressed on the WWII mobilization rate interacted with a 1960 year indicator variable, indicator variables of observation year, age, race, state of residence, state/country of birth. Fertility regressions also contain number of children older than 5. All indicator variables except state/country of birth and state of residence are also interacted with the 1960 year indicator variable. Instrumental variables used in regressions 4 and 5 are: 1940 male share ages 13-24 interacted with a 1960 year indicator variable, 1940 male share ages 25-34 interacted with a 1960 year indicator variable, and 1940 male share German interacted with a 1960 year indicator variable. All data is weighted using census person weights.

than 13 weeks per woman/year.

The results in Column 2 in Table 2 add individual-level controls for years of education as well as farm status. Introducing these controls lowers the size of the coefficient estimates, but the signs remain the same and the estimates for fertility, employment status, and marriage remain highly significant. Also adding marital status dummies further lowers the size of the estimates, but the effect on fertility remains quantitatively large, implying a mobilization effect of 0.27 on fertility at the average mobilization rate, which is a large fraction of the total increase in this measure of 0.37 between 1940 and 1960. Moreover, marriage, education, and farm status are all endogenous decisions that respond to some extent to the labor market changes implied by the war, so that it is not obvious whether these should be controlled for (this is particularly relevant for marriage, which is highly correlated with child bearing).

A potential concern about these regression results is that mobilization rates could be correlated with other state-level determinants of fertility, labor supply, and marriage that we do not control for. To address this possibility, in columns 4 and 5 in Table 2 we display results for 2-stage least squares (2SLS) regressions in which the state mobilization rates are instrumented using measures of the age structure of men and of German heritage.[11] More specifically, the instrumental variables are the share of males of age 13–24 among males of ages 13–44 in 1940 interacted with the 1960 dummy variable, the share of males of age 25–34 among males of 13–44 in 1940 interacted with the 1960 dummy, and the share of Germans among males age 13–44 in 1940 interacted with the 1960 dummy. The results of 2SLS regressions are similar to the previous estimates. The size of the estimates is even larger, especially in the case of fertility.

Table 3 presents analogous regression results for employment of the older age group 45-55. Women who were in this age group during the 1960 census were between 30 and 40 years old in 1945, and are thus likely to have entered the labor force during the war. The regression results confirm that mobilization had a sub-

[11] In the instrumental variable strategy we once again follow Acemoglu, Autor, and Lyle (2004), with the exception that rather than using the agriculture share as a state-level instrument, we control for farm status at the individual level (Acemoglu, Autor, and Lyle 2004 limited their sample to non-farm households).

Table 3: Impact of WWII Mobilization Rates on Labor Supply of Women Aged 45–55 (Coefficient Estimates from OLS and 2SLS Regressions for Variable "Mobilization Rate × 1960")

Dependent Variable	Regression				
	(1)	(2)	(3)	(4)	(5)
	Age 45–55 (N = 191715)				
Employed	0.058	0.185	0.174	0.604	0.522
	(0.074)	(0.081)	(0.077)	(0.221)	(0.180)
R^2	0.081	0.112	0.200	0.112	0.200
Weeks Worked	14.754	14.837	14.101	38.211	34.209
	(4.710)	(4.660)	(5.010)	(11.258)	(9.810)
R^2	0.054	0.074	0.154	0.074	0.154
p-value First Stage				<0.0001	<0.0001
Education and Farm Controls	no	yes	yes	yes	yes
Marital Status Controls	no	no	yes	no	yes

Notes: Standard errors (in parentheses) are adjusted for clusters of state of residence and year of observation. Estimates are from separate regressions of pooled micro data from 1940 and 1960 census. Regressions 1-3 are OLS and 4-5 are 2SLS. Each outcome variable is regressed on the WWII mobilization rate interacted with a 1960 year indicator variable, indicator variables of observation year, age, race, state of residence, state/country of birth. All indicator variables except state/country of birth and state of residence are also interacted with the 1960 year indicator variable. Instrumental variables used in regressions 4 and 5 are: 1940 male share ages 13-24 interacted with a 1960 year indicator variable, 1940 male share ages 25-34 interacted with a 1960 year indicator variable, and 1940 male share German interacted with a 1960 year indicator variable. All data is weighted using census person weights.

stantial long-run impact of these women's labor supply. The coefficient estimates on weeks worked are all highly statistically significant and imply an increase of employment by about seven weeks per year at the average mobilization rate, with even larger effects in the 2SLS regressions. The effect on the probability of employment is significant as long as education and farm status are controlled for.

Another potential concern about the empirical results is that employment and marriage are binary variables and fertility is a discrete variable. In Appendix A.1, we present regression results for probit regressions for employment and marriage and ordered probit regressions for the number of children under 5. Both qualitatively and quantitatively these results are very similar to our findings in Tables 2 and 3.

To summarize, our empirical results show that wartime mobilization is associated with higher fertility and lower labor force participation by young women during the baby-boom period. These findings do not yet pin down which mechanism provides the link between the war and young women's decisions in the following decades. In the next section, we develop a model that spells out such a link and assess its ability to account for the baby boom and baby bust.

3 The Model Economy

We now describe the model economy that we employ to quantitatively evaluate our explanation for the U.S. baby boom. At the aggregate level, the model is a version of the standard neoclassical growth model that underlies much of the applied literature in macroeconomics. We enrich this framework in three dimensions. First, we model married couples' life-cycle decisions on fertility and female labor-force participation. Since fertility and labor-force participation decisions are discrete, the model incorporates preference heterogeneity so as to generate heterogeneous behavior in these dimensions. Second, the production technology features limited substitutability of male and female labor, which implies that changes in the relative labor supply of men and women affect the gender wage gap. Third, we introduce a government that buys goods, employs soldiers, levies taxes, and issues debt. Modeling the government in detail will allow us to trace out the effects of war finance on labor supply and fertility.

3.1 Couples' Life-Cycle Choices of Fertility and Labor Supply

The model economy is populated by married couples who live for $T+1$ adult periods, indexed from 0 to T. Men work continuously until model period R, after which they retire. Women can choose in every period whether or not to participate in the labor market. Working women also retire after period R. Apart from deciding on labor supply, the main decision facing our couples is the choice of their number of children. Parents raise their children for I periods, at which time the children turn adult. All decisions are taken jointly by husbands and wives. A couple turning adult in period t maximizes the expected utility function:

$$U_t = E_t \left\{ \sum_{j=0}^{T} \beta^j \left[\log(c_{t,j}) + \sigma_x \log(x_{t,j} + x_{t,j}^W) \right] + \sigma_n \log(n_t) \right\}.$$

Here $c_{t,j}$ is consumption at age j of a household who turned adult in period t, $x_{t,j}$ is female leisure, and n_t is the number of children. Male leisure does not appear in the utility function, as men are continuously employed until retirement and their leisure is therefore fixed. $x_{t,j}^W$ represents a preference shock that shifts leisure preferences during war time. This shock can be interpreted to represent patriotism and allows us to match labor supply during the war. In regular times we have $x_{t,j}^W = 0$; what happens during the war is discussed in Section 4.1 below.

The before-tax labor income of a couple at age j who turned adult in period t is given by:

$$I_{t,j} = w_{t+j}^m e_{t,j}^m + w_{t+j}^f e_{t,j}^f l_{t,j}.$$

Here w_{t+j}^m is the male wage, $e_{t,j}^m$ is male labor-market experience (i.e., labor supply in efficiency units), w_{t+j}^f is the female wage, $e_{t,j}^f$ is female labor-market experience, and $l_{t,j}$ is female labor supply, which can be either zero or one (male labor supply is always assumed to be one). The flow budget constraint that a couple turning adult in period t faces in period $t+j$ is:

$$c_{t,j} + a_{t,j+1} = (1 + r_{t+j}) a_{t,j} + I_{t,j} - T_{t+j}(I_{t,j}, r_{t+j} a_{t,j}).$$

Here $a_{t,j}$ are assets (savings), r_{t+j} is the interest rate in period $t+j$, and $T_{t+j}(\cdot)$ is the income tax as a function of pre-tax labor and capital income. People are born and die without assets ($a_{t,0} = a_{t,T+1} = 0$).

For $j < R$ (i.e., until retirement age), labor market experience evolves according to:

$$\begin{aligned} e^m_{t,j+1} &= (1 + \eta_{m,j}) e^m_{t,j}, \\ e^f_{t,j+1} &= (1 + \eta_{f,j} l_{t,j} + \nu(1 - l_{t,j})) e^f_{t,j}, \end{aligned}$$

where $\eta_{m,j}$ is the age-dependent return to male experience, $\eta_{f,j}$ is the age-dependent return to female experience, and ν is the return to age for a woman who is currently not working.[12] We do not separately model the male return to age since men are continuously employed. Initial experience is normalized to one for both sexes, $e^m_{t,0} = e^f_{t,0} = 1$. For $j > R$ we have $e^m_{t,j} = e^f_{t,j} = 0$, i.e., men and women are no longer productive once they reach retirement age.

For young women who haven't had children yet, leisure is given by:

$$x_{t,j} = h - l_{t,j} - z_{t,j}.$$

Here h is the time endowment, and the variable $z_{t,j} \in \{0, \bar{z}\}$ is an adjustment cost (in terms of time) that has to be paid when a woman reenters the labor force, i.e., switches from non-employment to employment. This cost captures the job search effort as well as any other costs, pecuniary or emotional, that are incurred when reentering the labor force. The cost has to be paid only once for a female employment spell.[13] The general leisure constraint, which also includes the costs of having children, is given by:

$$x_{t,j} = h - \phi(n^y_{t,j})^\psi - \kappa\, b_{t,j} - l_{t,j} - z_{t,j}.$$

Here $n^y_{t,j}$ is the number of young (i.e., non-adult) children who are still living with their parents, $\phi > 0$ and $\psi > 0$ are parameters governing the level and curvature

[12] In principle, ν could be negative, implying that experience depreciates when women don't work. See Olivetti (2006) for a macroeconomic study of the importance of the return to experience for explaining female labor supply.

[13] More precisely, we have $z_{t,j} = \bar{z}$ if $l_{t,j} = 1$ and $l_{t,j-1} = 0$, and $z_{t,j} = 0$ otherwise. The main function of the reentry cost is to make female labor supply persistent, which is necessary to match the rise in female employment after the demand shock of World war II.

of the cost of children, $b_{t,j} \in \{0,1\}$ indicates whether a birth has taken place in period j, and $\kappa \geq 0$ is the additional time cost for the birth over and above the general time cost of children. Children live with their parents for I periods, after which they turn adult, form their own households, and are no longer costly to their parents. For simplicity, and realistically for the period, we assume that women who give birth do not work during the same period. Women can give birth only until age M, and only one birth per period is possible. Thus, for example, a woman planning to have three children must start giving birth at age $M-2$. The total number of children is given by the total number of births:

$$n_t = \sum_{j=0}^{M} b_{t,j},$$

and the number of non-adult children in any period is given by:

$$n_{t,j}^{y} = \begin{cases} \sum_{k=\max\{0,j-(I-1)\}}^{\min\{j,M\}} b_{t,k} & \text{if} \quad j \leq M+I-1, \\ 0 & \text{if} \quad j > M+I-1. \end{cases}$$

The population is heterogeneous in terms of the appreciation of leisure, i.e., the parameter σ_x in the utility function varies across couples.[14] In particular, in any cohort the distribution of σ_x is governed by the distribution function $F(\sigma_x)$. The distribution of σ_x determines the average female labor force participation rate at different ages. In the model, it is optimal for women to have children as late as possible, i.e., women work initially and then have children until they reach age M. Subsequently, only women with a relatively low appreciation for leisure (i.e., a low disutility for work) return to the labor force.

[14]Our aim is to introduce heterogeneous behavior in terms of fertility and labor-force participation while keeping the model parsimonious. We therefore introduce preference heterogeneity only along the leisure dimension. Introducing heterogeneity in terms of the preference for children would be less effective, because conditional on the number of children all women would have identical preferences at the labor-leisure margin. In principle, heterogeneous behavior could also arise with homogeneous preferences, as long as couples are indifferent between all bundles that are chosen in equilibrium. However, such a model would have the unattractive feature that aggregates are infinitely elastic with respect to infinitesimal price changes.

3.2 The Aggregate Production Function

The production technology is given by:

$$Y_t = A_t^{1-\alpha} K_t^{\alpha} \left(\theta (L_t^f)^{\rho} + (1-\theta)(L_t^m)^{\rho}\right)^{\frac{1-\alpha}{\rho}},$$

where A_t is productivity, K_t is the aggregate capital stock, L_t^f is female labor supply in efficiency units, and L_t^m is male labor supply in efficiency units. The aggregate capital stock depreciates at rate δ per period. The production function allows for limited substitutability between male and female labor, governed by the parameter ρ. Productivity increases at a constant rate γ every period:

$$A_{t+1} = (1+\gamma)A_t.$$

In the balanced growth path, the growth rate of output per capita will be equal to γ. The production technology is operated by perfectly competitive firms, so that all factors are paid their marginal products and profits are equal to zero in equilibrium.

3.3 The Government

The government buys goods G_t (which include military goods), drafts soldiers L_t^D into the military (measured in efficiency units of male labor), and finances government spending via taxes and government debt B_t. We consider a tax system consisting of a flat capital income tax τ_k, a flat labor income tax τ_l with an exemption level of of ξ_t, and a lump-sum tax τ_{LS}, so that the tax function is:

$$T_t(I_l, I_k) = \tau_{l,t} \max\{I_l - \xi_t, 0\} + \tau_{k,t} I_k + \tau_{LS,t},$$

where I_l is labor income and I_k is capital income. Following Ohanian (1997) and McGrattan and Ohanian (2010), we assume that the monetary compensation of draftees equals the wage received by comparable civilian workers. This formulation has the advantage that we do not need to distinguish between draftees and civilians in the formulation of the household problem. Let P_t denote the size of cohort t (i.e., the number of couples who enter adulthood in t). The government budget constraint is:

$$G_t + w_t^m L_t^D + (1 + r_t)B_t = B_{t+1} + \sum_{s=1}^{T} P_{t-s} \int_0^\infty T_t(I_{t-s,s}, r_t\, a_{t-s,s})\, dF(\sigma_x),$$

The government budget constraint shows how government spending $G_t + w_t^m L_t^D$ and service of existing debt $(1 + r_t)B_t$ is financed through issuing new debt B_{t+1} as well as through tax revenue.

3.4 Market Clearing

The market-clearing condition for capital is given by:

$$K_t + B_t = \sum_{s=1}^{T} P_{t-s} \int_0^\infty a_{t-s,s}\, dF(\sigma_x),$$

that is, the sum of the capital stock and government debt is equal to the sum of the assets of all cohorts that are currently alive, where in the integral it is understood that assets $a_{t-s,s}$ are a function a household's leisure-preference parameter σ_x. Similarly, the market-clearing condition for male labor is given by:

$$L_t^m + L_t^D = \sum_{s=0}^{R} P_{t-s} \int_0^\infty e_{t-s,s}^m\, dF(\sigma_x),$$

and female labor supply satisfies:

$$L_t^f = \sum_{s=0}^{R} P_{t-s} \int_0^\infty e_{t-s,s}^f\, l_{t-s,s}\, dF(\sigma_x).$$

Here L_t^m and L_t^f are the total efficiency units of labor that men and women supply to the civilian labor market. Finally, given that children turn adult after spending I periods with their parents, the cohort sizes P_t evolve according to the law of motion:

$$P_{t+I} = \frac{1}{2} \sum_{s=0}^{M} P_{t-s} \int_0^\infty b_{t-s,s}\, dF(\sigma_x).$$

The factor $\frac{1}{2}$ enters the law of motion because fertility is measured in terms of individuals while cohort size is measured in terms of couples. More precisely, $b_{t-s,s}$ describes the number of births (zero or one) of a couple born in period $t-s$ at time t. Integrating over all these couples and multiplying by cohort size P_{t-s} gives the total number of children born in period t to parents from cohort $t-s$. Summing this over all cohorts who are in childbearing age in period t (i.e., those aged zero to M) yields the total number of children born in period t. Dividing by 2 results in the number of couples P_{t+I} turning adult I periods later.

3.5 What Drives Fertility?

Before turning to quantitative results, it is instructive to consider how fertility and female labor-force participation decisions are determined in the model. In the calibration considered below (Section 4), all women initially enter the labor force when turning adult, and then quit in order to have children. It turns out to be optimal to have children as late as possible, because then the initial earnings period can be extended and the time cost of having children can be delayed. Women therefore have children right up to the final fecund period M. A key implication of this timing of fertility is that the marginal child is the first one: women who want to have an additional child must leave the labor force one period earlier. What then determines whether a woman will have an additional child?

Consider, first, the case of a woman who does not anticipate reentering the labor force after having children. For such a woman, both the marginal utility of having another child and the disutility (in terms of reduced leisure) of raising the child are fixed numbers. The only variable part of the tradeoff is the opportunity cost of having to exit the labor force earlier, which depends on forgone wage income in this period. Thus, young women's wages are a key determinant of fertility. However, what matters is not the absolute level of the young female wage, but the product of the wage and the marginal utility of consumption. The marginal utility of consumption, in turn, is driven by the present value of a couple's lifetime income. Given that the remainder of household income is earned by husbands, fertility ends up being determined by female wages relative to male wages. In a balanced growth path, female wages increase in proportion to male wages, so that fertility rates are constant. During the transition after the war shock, in contrast, relative wages will fluctuate, leading to changes in fertility.

The tradeoff for having another child is more complicated for women who would adjust their labor supply later in life if they had another child, in which case relative wage at older ages is also relevant. However, this margin operates only for relatively few women. The fertility implications of the war shock in our model are therefore primarily driven by the shock's impact on young women's wages relative to young couples' lifetime income. A second channel through which the war affects fertility is changes in labor taxation. To assess the quantitative significance of these channels, we now turn to the calibration procedure for our model economy.

4 The Quantitative Experiment: World War II and the Baby Boom

We now would like to assess the impact of the shock of World War II on subsequent fertility. We first discuss how the war is modeled as both a shock to the labor market and as a shock to government spending and taxes. Next, since we are looking for quantitative results, we calibrate the model economy to match certain characteristics of the United States in the pre- and post-war periods as well as during the war itself. We then present our main results and discuss the sensitivity of the findings to alternative parameterizations and assumptions.

4.1 Modeling the War Shock

Our overall computational strategy is to model World War II as an unexpected shock that displaces the economy from a pre-war balanced growth path. The representation of the war builds on the analyses of Ohanian (1997), Siu (2008), and McGrattan and Ohanian (2010) of the fiscal implications of the war in a neoclassical framework. The war shock consists of three components. First, during the war the government drafts men for military service. We set the number of draftees to $L_t^D = 0$ both before and after the war and to a positive number $L_t^D > 0$ during the war. Draftees are not available for civilian production, so that the male labor input in the production function drops during the war.

Illustration 1: U.S. Government World War II Recruitment Posters

The second aspect of the war shock is a change in fiscal policy. The war led to a massive increase in government spending, which was financed by higher taxes and a large increase in government debt. Accounting for the fiscal implications of the war is important for our analysis, because our theory revolves around work incentives for young women during the baby boom period, which are affected by marginal labor taxes. Accordingly, we match the increase in government spending during the war to data and allow labor taxes to increase permanently. Government debt also increased during the war. For the purposes of our analysis, government debt matters because of the fiscal burden that it places on families during the baby boom period. With this in mind, we set the level of government debt at the end of the war such that ratio of debt service to GDP in the post-war period matches the data.

The third component of the war shock consists of a "patriotism" shock that increases female labor force participation during the war. In particular, the preference shock $x^W_{t,j}$ is set to zero both before and after the war. In contrast, we set $x^W_{t,j} = \bar{x}^W > 0$ for those women who enter the labor force during the war. $\bar{x}^W$ is chosen such that the rise in the overall female labor-force participation rate during the war matches the data. Given that our theory is about how the female labor market after the war is affected by the rise of female employment during the war, matching this increase is essential for our exercise.

In principle, one might expect that female participation should rise even without a patriotism shock, because of the drop in male labor supply and also the wealth effect of high taxes. Indeed, in the analysis of McGrattan and Ohanian (2010) these factors are sufficient to explain the rise in female employment. However, McGrattan and Ohanian use a model with infinitely-lived agents and focus exclusively on the war period, whereas we employ a life cycle model with a parametrization that is geared towards being consistent with evidence on female labor supply and fertility from before the war. We find that our model does not reproduce the large wealth effect on labor supply that drives the results in McGrattan and Ohanian. Rather, our findings line up with the analysis of Mulligan (1998), who argues that in the United States after-tax real wages actually fell during the war, so that other factors, such as patriotism, are required to explain the large increase in female labor-force participation. Indeed, the U.S. government ran a public campaign to recruit women for the war effort (see Illustration 1). Whereas previously society was often prejudiced against the employment of married women[15], during the war joining the labor force was actively encouraged. The patriotism shock captures this change.

4.2 Calibrating the Model Parameters

We now describe in detail how all model parameters as well as the different components of the war shock are chosen. Given that the war shock includes a permanent change in tax rates, after the war the economy converges to a new balanced growth path corresponding to the new fiscal environment. We calibrate the model such that the pre- and post-war balanced growth paths matches a specific set of characteristics of the actual U.S. economy. For the macro side of the model, we choose a set of target moments that characterize long-run U.S. growth and that are standard in the real-business-cycle literature. Fertility and patterns of female labor-force participation are matched to observations in the pre-war

[15] See for example Goldin (1990) for a discussion of marriage bars (which excluded married women from employment in certain professions, in particular clerical work and teaching), wich were widely practiced before World War II.

period, mostly from the 1940 census. The war shock is calibrated to match data on mobilization rates, female labor-force participation, and fiscal changes during World War II.

One central aspect of the calibration is to pin down how strongly fertility and labor supply react to the war shock. Here our strategy is to constrain the model such that it is consistent with the cross-state evidence on the impact of mobilization rates on fertility presented in Section 2. Relying on this evidence is ideal for our purposes, because the empirical setting consists precisely of the historical episode that we are trying to understand. Of course, this strategy implies that the quantitative results from the model do not provide independent evidence on the magnitude of the reaction of fertility to the war shock. Rather, the added value of the quantitative analysis is to make explicit the causal connection between the war and the rise in fertility, to assess the implications of the theory for changes in female labor supply (which are not constrained by the calibration) as well as the timing of the rise and fall in fertility, and to enable us to carry out counterfactual experiments that highlight the relative importance of alternative channels linking World War II and the baby boom.

The first calibration choice concerns the length of a model period. The main characteristic that defines a period in the model is that women can have one child per period. In the balanced growth path, once they start to have children women give birth to a child every period until reaching the fecundity limit M. The length of the model period therefore corresponds to the average time between births. In the United States, the average spacing of births narrowed from over three years for the cohort of mothers born 1916–1920 to slightly above two years for the cohort 1931–35 (Whelpton 1964). As a compromise, we set the model period as corresponding to 2.5 years in the data. We also set the length of childhood to $I = 8$, so that the age of adulthood corresponds to 20 years in the data. The fecundity limit is set at $M = 4$ (women are fecund until 32.5 years old), the last period of work is $R = 15$ (retirement starts at age 60), and the last period of life is $T = 19$ (people die at age 70). In the real world, of course, women can conceive at ages older than 32.5, but the likelihood of conception declines from the early 30s. More importantly, in the model women have children right up to the end of their fertile period, which implies that the fecundity limit determines the average age at first birth. Choosing a higher fecundity limit would imply a counterfactually high age of first-time mothers.

At the aggregate level, we match the capital income share, the depreciation rate, and the return to capital to long-run U.S. data. Where possible, we use the same calibration targets as Greenwood, Seshadri, and Vandenbroucke (2005) for these macroeconomic statistics to yield comparable results. Consequently, we set the capital income share to 0.3 ($\alpha = 0.3$), the depreciation rate to 4.7 percent per year ($\delta = 1 - (1 - 0.047)^{2.5}$), and the annualized pre-tax return to capital to 6.9 percent.[16] The return to capital is a function of the capital-output ratio, which, in turn, is mostly governed by the time-preference parameter β. Given the other calibration choices, the return to capital is matched by setting $\beta = 0.987$.[17] We also follow the real-business-cycle literature in assuming that full-time work takes up one-third of discretionary time (Cooley and Prescott 1995). Given that the time cost of full-time work is normalized to one (i.e., $l_{t,j} \in \{0, 1\}$), the time endowment is set to $h = 3$. The parameters ρ and θ govern the substitutability between male and female labor, as well as relative wages. The share parameter θ is chosen to match a ratio of average female to male wages of 0.66 in 1940, which results in $\theta = 0.35$.[18] The elasticity parameter ρ has been estimated by Acemoglu, Autor, and Lyle (2004) using census data. They suggest a range of 0.583 to 0.762 for ρ; following this estimate, we set $\rho = 0.65$. The implied elasticity of substitution between male and female labor is about 2.9. The annualized productivity growth rate of the economy is set to 1.8 percent ($\gamma = 1.018^{2.5}$), which corresponds to the average growth rate of real GDP per capita in the U.S. during the period 1950-2003.[19]

The parameters governing the returns to experience in the labor market deter-

[16] See Cooley and Prescott 1995 for details on how these statistics can be computed from aggregate U.S. data. The return to capital is matched for the post-war steady state; however, matching the return in the pre-war steady state leads to mostly identical results.

[17] Since we model a life-cycle economy, the direct correspondence between the discount factor, the growth rate, and the return to capital that holds in infinitely-lived agent economies does not apply in our framework.

[18] Average wages are computed across ages 20–60 and all race groups from the 1940 Census, see Appendix A.3 for details.

[19] Data from Penn World Table Mk. 6.2, see Heston, Summers, and Aten (2006).

mine the steepness of age-wage profiles, both in the cross section and over the life cycle. To calibrate the experience accumulation function, we estimate an earnings equation for men using data from the 1940 census. The earnings equation contains linear and quadratic terms in experience, and we choose the $\eta_{m,j}$ parameters to match the empirical estimate of the return to experience. The resulting parameter values are given by:

$$\eta_{m,j} = \exp(0.125 - 0.00053(12.5j - 6.25)) - 1.$$

We also assume that the return to experience for women and men is the same, $\eta_{f,j} = \eta_{m,j}$ for all j. We then choose the return to age ν such that in the pre-war balanced growth path, at age 32.5 (when women reach the end of the fecund period) the productivity of working women is larger by a factor of 1.42 than at age 20. This factor is obtained by estimating an earnings function for women and predicting women's wages at ages 20 and 32.5. The return parameter matches this ratio by setting the slowdown in experience accumulation when women leave the labor force for childbearing. The procedure yields a return to age (per model period) of $\nu = 0.003$ (thus, the return to age is close to zero).[20]

The child cost parameters are chosen such that the average private cost of a child (which in the model consists of forgone female earnings) amounts to 40 percent of GDP per capita in the balanced growth path, thus matching the estimate by Haveman and Wolfe (1995) of the total private cost of a child in the United States.[21] The curvature parameter in the child cost function ψ (which determines the returns to scale to having children) and the additional cost of young children κ are estimated from U.S. time-use data, which results in $\psi = 0.33$ and $\kappa = 0.209$ (see Appendix A.5 for details). Given these choices, the overall cost of children is matched to its target by setting the level parameter of the child cost function to $\phi = 0.412$.

Turning to preferences, we impose a uniform distribution for the taste for leisure σ_x in the population. The distribution of σ_x together with the fertility weight in

[20] Further details on the calibration of the accumulation of experience are provided in Appendix A.4.

[21] This approach to calibrating the cost of children has been previously used by Doepke (2004) and Lagerlöf (2006), among others.

utility σ_n determine female labor force participation, the level of fertility, and the elasticity of the fertility reaction to the war shock. Intuitively, fertility decisions are discrete and take place on the extensive margin. Women with different numbers of children are distinguished by their σ_x, and the density of the distribution for σ_x around the cutoffs between women with different numbers of children determines how elastic fertility is in the aggregate. We choose these parameters to match three targets. First, we set the labor-force participation rate of married women aged 33–60 in the pre-war balanced growth path to 13 percent, the observed value for the United States in 1940.[22] Second, we target a completed fertility rate of 2.4 in the pre-war steady state, which matches the completed fertility rate of women born between 1911 and 1915, who were in their prime fertility years (average age 27) in 1940. We match a completed fertility rate rather than the total fertility rate because total fertility rates are sensitive to changes in the timing of births. Third, we choose the distribution of σ_x to match the empirical evidence in Section 2. Specifically, regression (2) in Table 2 (which includes education and farm controls) yields an estimate of 0.665 of the impact of the state mobilization rate on the number of children under age 5 in 1960. We compute an analogous statistic in our model by comparing the average number of children under age 5 in 1960 in two different scenarios, one in which we match the actual mobilization rate during the war, and a counterfactual simulation in which we set the mobilization rate during the war to zero, while keeping everything else the same.[23] The model-implied regression coefficient is the fertility difference between the war scenario and counterfactual scenario, divided by the mobilization rate. However, one concern about this procedure is that the number of children under age 5 is a measure of period fertility. We know that empirically, period fertility increased by much more than cohort fertility during the baby boom, due to a change in the timing of births. Given that the timing of births is fixed in our model, it is appropriate to target the smaller change in cohort fertility. We therefore divide the target for the reaction in fertility by the ratio of the increase in total fertility to cohort fertility during the baby boom, resulting in a more conser-

[22]Data from the 1940 U.S. Census, see Appendix A.3 for details.

[23]The counterfactual simulation still includes all fiscal changes related to the war. The fiscal changes do not vary at the state level and are therefore not picked up by cross-state regressions, so that the fiscal side needs to be held constant in the simulations as well.

vative parameterization with smaller changes in fertility.[24] This procedure yields upper and lower bounds for the distribution of leisure taste of $\min(\sigma_x) = 1.154$ and $\max(\sigma_x) = 1.612$.

Before the war, income taxes and government debt were low. From 1932 to 1940, the tax rate for the lowest bracket of the federal income tax was 4 percent, and the next-higher bracket started at more than triple average household income.[25] Consequently, we set the labor income tax to 4 percent in the pre-war steady state and set government debt to zero. During the war, taxes rapidly increased and remained high subsequently. The marginal tax rate for average-income households reached 22 percent in 1943, and moved between 20 and 25 percent from 1944 to 1964. We model this change as a permanent jump (starting from the war period) in the marginal tax on labor income to 22 percent, which is the average marginal tax at average income from 1943 to 1960. Unlike labor taxes, capital taxes were already fairly high before World War II. For simplicity, we set capital taxes to a constant level throughout. The level is chosen such that revenue from capital taxes matches the total revenue from the corporate income tax, a proportionate share of individual income tax (based on the share of capital income in total income), and federal excise taxes to GDP from 1950 to 1960.[26] This procedure results in a tax rate of $\tau_k = 0.45$.[27]

The increase in government spending during the war is partially financed through government debt that is repaid after the war. Government debt matters for our analysis through the fiscal burden it presents during the baby boom period. To this end, we set the level of government debt to match the amount of interest payments on government debt to data. From 1946 to 1960, interest on govern-

[24] Cohort fertility increased by 0.93 from 1910 to the peak in 1932, whereas the total fertility rate increased by 1.37 from 1940 to 1957. We therefore divide the target by $1.37/0.93 = 1.47$.

[25] For data on average household income we rely on Piketty and Saez (2007).

[26] An alternative to matching actual tax rates and tax revenue is to follow McGrattan and Ohanian (2010) and rely on the estimates of Joines (1981) of effective marginal taxes. Results would be broadly similar, with the main difference that Joines' estimates of marginal capital taxes are about 10 percentage points higher than our estimate for the post-war period. However, given our focus on fertility behavior, we are more interested in matching tax rates faced by average households as opposed to average owners of capital (who are rich), so that we use numbers that are not driven by tax rates in high income brackets faced by a small number of taxpayers.

[27] Data from "Budget of the United States Government, Fiscal Year 2012," Government Printing Office.

ment debt averaged 1.5 percent of GDP, with little variation from year to year. Consequently, we assume that after the war shock, the economy converges to a new balanced growth path with a constant debt/GDP ratio and a ratio of interest payments to GDP of 1.5 percent. During the transition to the balanced growth path, government debt is assumed to increase at the trend growth rate of total GDP from year to year. This procedure implies the amount of government debt outstanding at the end of the war.

We next turn to government spending G_t. For the pre-war balanced growth path, we set government spending to balance the government budget given tax revenue in the absence of government debt. During the war period, we set G_t to 44.5 percent of GDP, which corresponds to the average of government expenditures as a fraction of GDP for the period 1943 to 1945. To close the government budget constraint during the war period, we allow the government to levy a one-time lump-sum tax.[28] Government spending dropped rapidly right after the war, and remained around 20 percent of GDP until the 1970s. We assume that after the war the economy converges to a balanced growth path in which government spending is at 19.9 percent of GDP, which is the average spending/GDP ratio from 1950 to 1960 in the data. The exemption level ξ_t for labor income is set to zero for the pre-war balanced growth path, and for the post-war balanced growth path we choose ξ_t to balance the government budget given the other assumptions on taxes, spending, and debt.[29] This exemption level amounts to 17.5 percent of average income. During the transition to the balanced growth path, the exemption level grows at the trend growth rate. We adjust the level of government spending period-by-period during the transition to balance the government budget. The resulting fluctuations in spending are small, with the spending varying between 19.9 and and 20.9 percent of GDP.

The remaining elements of the war shock are the mobilization of male soldiers

[28] The lump sum tax is equal to zero in all other periods. An alternative way to close the government budget constraint (pursued by Siu 2008) is to allow for an even higher level of government debt, part of which is then removed through surprise inflation right after the war. Since a surprise inflation has similar effects to a lump-sum tax, we opted for the simpler modeling option.

[29] This is almost equivalent to the procedure in McGrattan and Ohanian (2010), who balance the budget with a lump-sum rebate, except that the exemption does not benefit retired households without labor income.

Parameter	Interpretation	Value
I	Duration of childhood	8
M	Final period of fecundity	4
R	Final period before retirement	15
T	Lifespan	19
α	Capital share	0.3
δ	Capital depreciation rate	0.113
θ	Weight of female labor in technology	0.35
ρ	Elasticity of substitution parameter	0.65
γ	Productivity growth rate	0.046
ν	Return to age	0.003
β	Time discount factor	0.987
σ_n	Utility weight on fertility	1.53
$\min(\sigma_x)$	Minimum utility weight on leisure	1.154
$\max(\sigma_x)$	Maximum utility weight on leisure	1.612
$\bar{z}$	Cost of labor market reentry	1.25
h	Time endowment	3
ϕ	Level of time cost of children	0.412
ψ	Curvature of time cost of children	0.33
κ	Additional cost of young children	0.209
τ_k	Marginal tax on capital income	0.45
$\tau_{l,pre-war}$	Pre-war marginal labor tax	0.04
$\tau_{l,post-war}$	Post-war marginal labor tax	0.22
$\bar{x}^W$	Patriotism shock	1.25

Table 4: Calibrated Parameter Values

and the patriotism shock. Given that a model corresponds to 2.5 years (see Section 4.2 below) and that mass mobilization did not start before 1943, we assume that the war lasts for a single period. During the war period, L_t^D is set to 30 percent of the total male labor force, which matches the actual male mobilization rate during the final years of the war.[30] Hence, the male labor input in the production function drops by 30 percent during the war. The patriotism preference shock $\bar{x}^W$ (which lowers the disutility of labor for women entering the labor force during the war) as well as the fixed cost $\bar{z}$ for reentering the labor market itself govern the persistence of female labor supply and the size of the increase in female labor for participation during the war. Given that the theory is about the long-term implications of the changes in female labor supply caused by the war, it is essential for our analysis to generate a persistent increase in female employment as a result of the war shock. We choose $\bar{x}^W$ to match an overall female labor-force participation rate of 34 percent in the war period (Acemoglu, Autor, and Lyle 2004), which yields $\bar{x}^W = 1.25$. Once $\bar{z}$ is set above a certain threshold, most women who enter the labor force during the war continue working after the war, given that the fixed cost for entry has already been paid. It turns out that different values for $\bar{z}$ above the threshold lead to very similar predictions (provided that σ_n and the distribution of σ_x are adjusted accordingly to match the targets for fertility and overall female labor supply). For simplicity we therefore set $\bar{z} = \bar{x}^W$. The calibrated parameter values are summarized in Table 4.

4.3 The Impact of the War on Post-War Fertility

Even though the war shock in the model lasts only for one period, it has long-term consequences. One reason for long-term effects is the rise in government debt during the war and the related shift to higher tax rates. The second reason for long-term effects is persistence in female labor supply. For the most part, the war draws older women into the labor force (the youngest women are working anyway, and women who are currently having children are less willing to enter). Once these women have paid the fixed cost of entering the labor market and have accumulated experience, they choose to stay on working after the war. This increases the ratio of female to male labor supply, and depresses female wages.

[30] See Appendix A.6 for details.

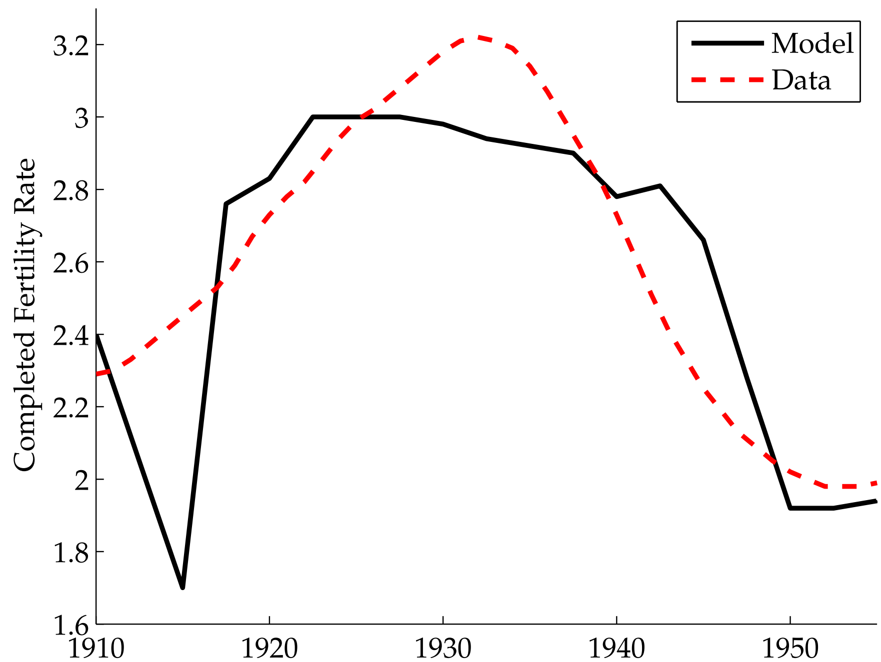

Figure 6: Completed Fertility Rate by Birth Cohort (i.e., by Birth Year of Mother), Model versus Data

It is this decline in the relative female wage after the war which is responsible for most of the long-term effects of the shock.

Figure 6 displays the response of the cohort fertility rate to the war shock in the model. Some women born in 1915 start to work during the war rather than having another child, resulting in low fertility for the cohort. Fertility rises for the following cohorts. The women born between about 1920 and 1940 begin childbearing after the war, while also facing increased labor-market competition from the older war-generation of women. Among these younger women, many decide to leave the labor market earlier in order to have another child, resulting in higher cohort fertility. In essence, the experienced war generation crowds out young women from the labor market. In the model simulation, cohort fertility peaks at 3.0, compared to a peak of 3.2 in the U.S. data. Thus, the model accounts for most of the increase in cohort fertility during the baby boom period. The model also accounts for the baby bust. By the time women born in the mid to

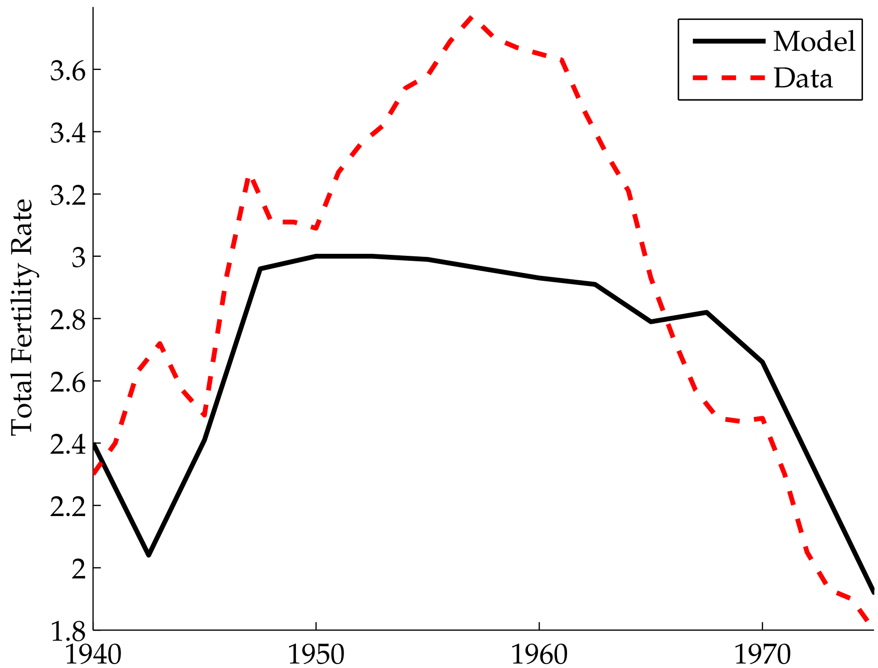

Figure 7: Total Fertility Rate, Model versus Data

late 1940s enter the labor market, the war generation of women has mostly retired, which relieves the pressure on the female labor market. As a consequence, these women work more and have fewer children, resulting in a drop in cohort fertility that closely matches the data.

As Figure 7 shows, the model explains a comparatively smaller fraction of the increase in the total fertility rate, which rises to a maximum of 3.8 in the data. The deviation between the patterns for total fertility and cohort fertility are due to changes in the timing of births. In particular, in the mid-1950s older cohorts of women were having children late while younger women were having them early, resulting in total fertility rates much higher than the lifetime fertility rates of any given cohort. Given that our model does not capture these shifts in the timing of births, it cannot account for the larger increase in the total fertility rate. Nevertheless, the model is still able to generate a sizeable increase in total fertility and matches the timing of the rise and fall of fertility rather well.

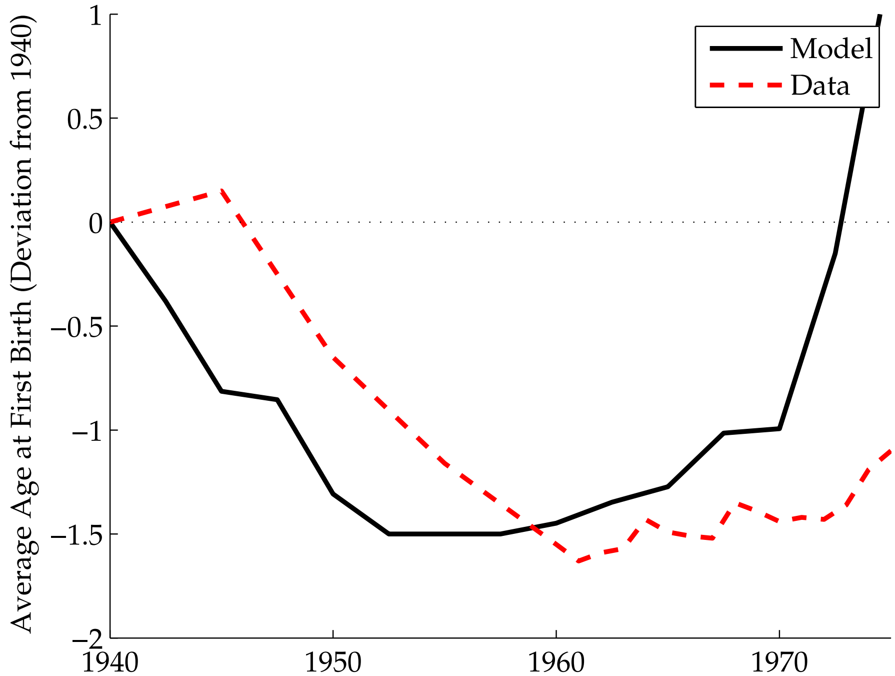

Figure 8: Average Age at First Birth, Deviation from 1940, Model versus Data

The increase in fertility predicted by our model is generated by a specific mechanism: younger women who did not work during the war are crowded out of the labor market, and decide to start having children at a younger age. We can evaluate this mechanism by comparing changes in the average age at first birth between model and data, which is done in Figure 8. In both model and data, the average age at first birth drops substantially during the baby boom period, and then recovers. The initial decline proceeds slightly more quickly in the model, but the size of the reduction in the age at first birth is almost identical between model and data. These results suggest that mechanism for fertility reduction captured by the model is indeed empirically relevant.

4.4 Implications for the Female Labor Market

In our theory, the increase in fertility during the baby boom is driven by changes in the female labor market. It is therefore important to check that the model

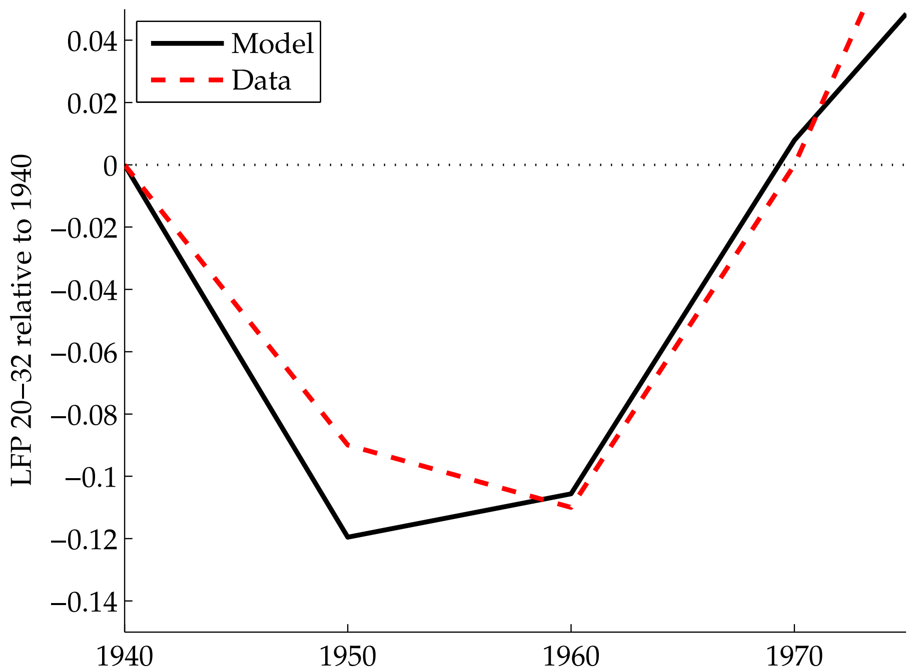

Figure 9: Average Labor Force Participation of Young Women, Deviation from 1940, Model versus Data. Data for ages 20–32 is for all women. See Appendix A.3 for details.

matches the data in this dimension as well. Figure 9 displays labor-force participation rates for women aged 20–32 (the prime child-bearing years) throughout the transition after the war shock. Even though changes in labor supply are not pinned down by the calibration procedure, the model matches the data remarkably well. For 1950, the decline in young female labor supply is slightly larger in the model compared to the data, and for 1960 and 1970 the match is almost exact.

In the model, the changes in labor supply and fertility are ultimately driven by changes in the wages paid to young women. For the results to be plausible, it is important that the model does not overstate the wage implications of the rise of female employment. Figure 10 displays the average wage of young women (20–24) in the model as a fraction of the average wage of men in the same age group. Since these women are permanently employed, there is no variation in average labor market experience in this group over time, so that variations in the wage

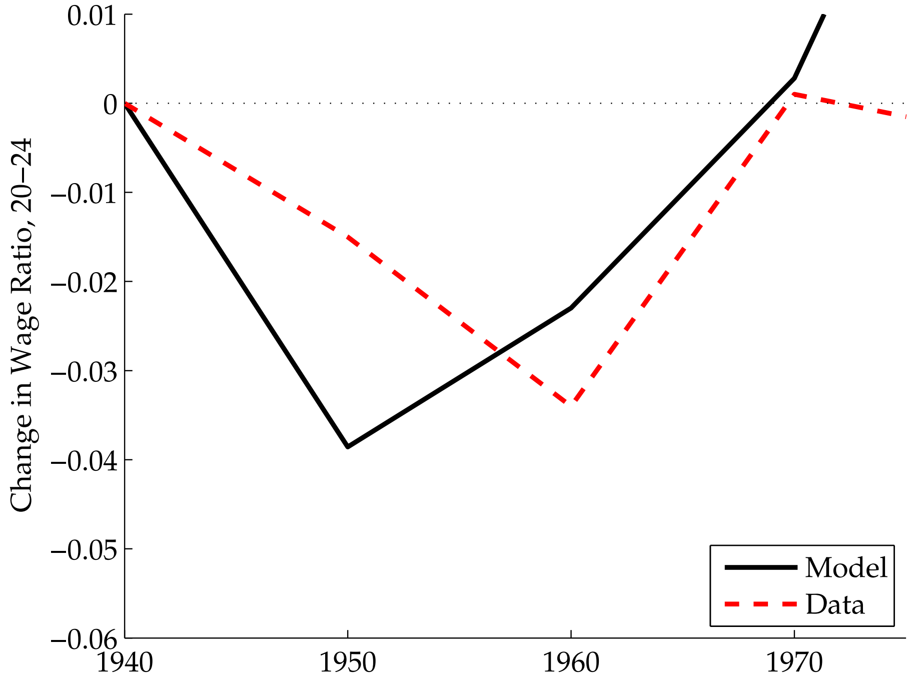

Figure 10: Change in Ratio of Average Female to Average Male Wage for Single Women aged 20–24, Model versus Data. See Appendix A.3 for details.

are entirely due to changes in the wage per efficiency unit of female labor. The change in relative female wages in the model is close to what is observed in the data.[31] However, the timing is somewhat different, with the model generating a larger drop in relative wages in 1950 compared to 1960, with the opposite pattern being observed in the data.[32]

One potential explanation for the different timing in relative wage changes is that in the real world additional factors were present that moved relative female wages, such as gender-biased technological change. Our theory suggest

[31]Notice that female wages decline only in relative, but not in absolute terms: sustained productivity growth implies that average wages rise for both men and women.

[32]It may be the case that the data understate the true decline in the relative female efficiency wage between 1940 and 1950, because the average education of women increased relative to men during this period. Selection issues may also be present, but are unlikely to be a major problem because (in the data) we focus on the wages of single men and women aged 20–24, the vast majority of whom were working.

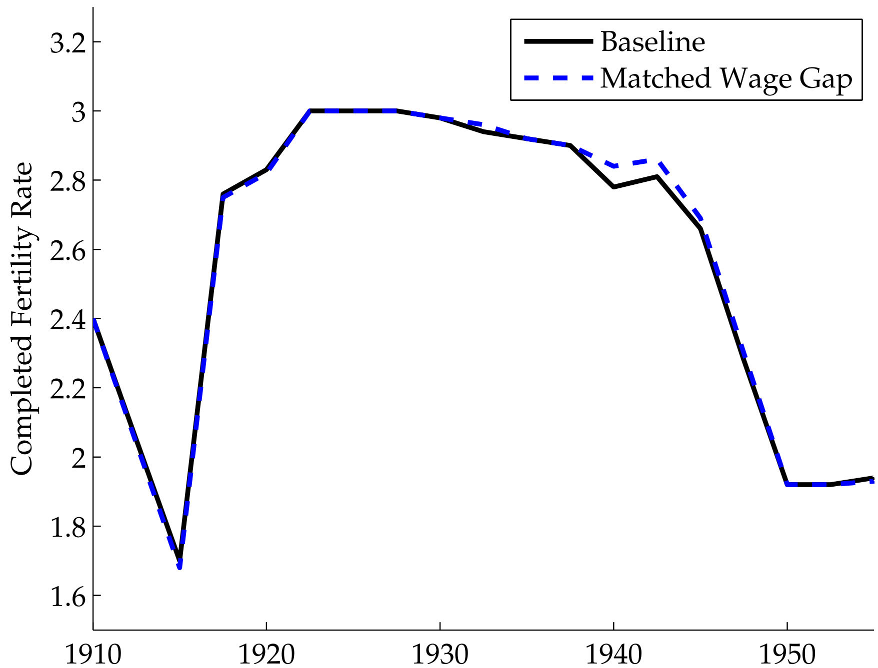

Figure 11: Completed Fertility Rate by Birth Cohort (i.e., by Birth Year of Mother), Baseline versus Matched Wages

that such other factors should also have an impact on fertility rates. As a consequence, one may wonder how well our predictions for fertility would hold up if other sources of shifting relative wages were incorporated in the theory. Figure 11 answers this question. We compare the fertility predictions of the baseline model with an alternative model in which we allow θ, the relative weight of female labor in the production function, to vary over time. Specifically, we choose the time path for θ such the changes in relative female wages from 1940 to 1950, 1960, and 1970 are matched exactly (i.e., in Figure 10 model and data would coincide), with a linearly interpolated path within each decade. Perhaps surprisingly, the fertility predictions of the baseline case and the matched-wages simulation are almost identical, with the main difference being a slightly higher fertility rate in the matched-wages case for cohorts who have their babies around 1960. This difference arises because in the data relative female wages in 1960 are lower compared to our baseline simulation. In principle, one would expect that the higher relative female wages in 1950 would lead to lower fertility compared to the baseline for earlier cohorts. However, this effect remains small due to the discrete nature of fertility. Most women in these cohorts are having three children, and only very few of them are close to being indifferent between two and three. While the higher relative wages in the matched-wages simulation generally imply lower fertility, for most women the difference is not large enough to change their decision.

4.5 The Labor Supply Channel versus the Fiscal Channel

In our analysis there are two principal channels through which the war affects fertility in the post-war period. First, there is the labor market channel, i.e., the war generation of women continues to work after the war, which creates pressure on the female labor market. The second channel is the fiscal channel: War expenditures lead to the issuance of government debt and higher taxes in the post-war period. Notice that the empirical results in Section 2 speak only to channels that vary at the state level, which does not include fertility changes that are due to higher federal income taxes.

We can use our model to decompose the changes in fertility into a fiscal channel and a labor-market channel. To this end, Figure 12 compares the fertility predictions of the baseline simulation to a counterfactual simulation that removes all fiscal changes. More precisely, in this alternative simulation there is no change in taxes, no rise in general government spending, and no issuance of government debt during or after the war. However, the labor market consequences of the war (mobilization and the patriotism shock that increase female labor supply) are still present. Given that in this simulation the war does not have permanent effects, the economy ultimately converges back to the pre-war balanced growth path. We can interpret the results for this simulation as the implications of the labor market channel alone, whereas the gap between the baseline simulation and the simulation without fiscal shocks is what is generated by the fiscal consequences of the war.

The results suggest that both the labor market channel and the fiscal channel contribute to the size of the baby boom, but that the contribution of the labor

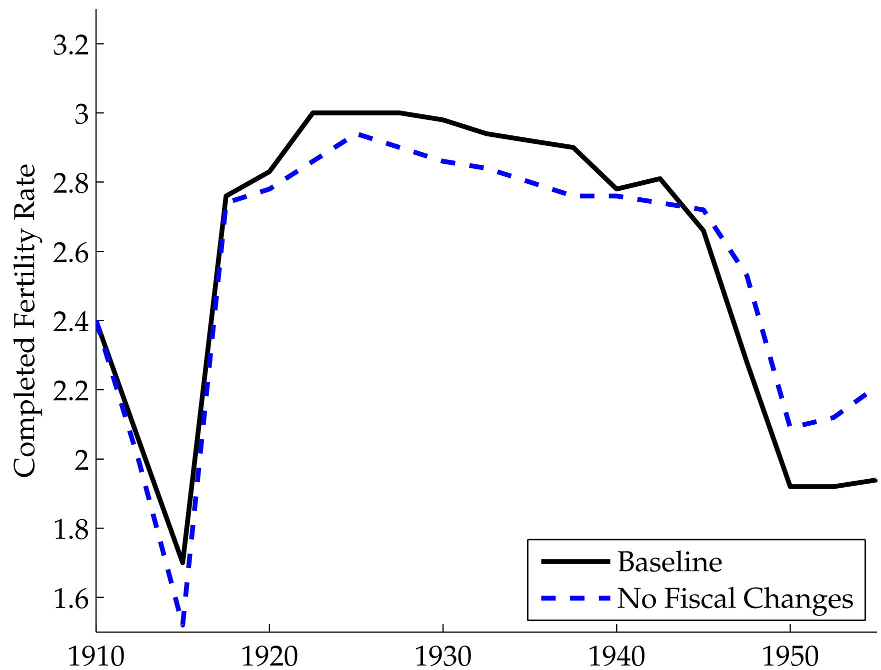

Figure 12: Completed Fertility Rate by Birth Cohort (i.e., by Birth Year of Mother), Baseline versus No Fiscal Changes

market channel is quantitatively much more important. During the height of the baby boom, fertility is higher by about 0.1 in the baseline simulation (with both the labor market channel and the fiscal channel) compared to the simulation with only the labor market channel. Relative to a pre-shock fertility rate of 2.4, the labor market channel accounts for about 80 percent of the maximum rise in fertility, with the remaining 20 percent due to the fiscal channel. The intuition for this finding is that the rise in taxation after the generates offsetting income and substitution effects that are of similar size. In fact, given that the model is consistent with balanced growth, a proportional income tax would generate exactly offsetting effects and have no impact on fertility at all. The reason why we are getting at least some impact on fertility is that the exemption level for labor income drives a wedge between the marginal and the average tax rate. However, given that marginal tax rates were still fairly moderate, this wedge does not lead to a large reaction in fertility.

5 International Evidence: Allied versus Neutral Countries in World War II

We now turn to international evidence to assess the empirical relevance of our mechanism from a different perspective. Most industrialized countries experienced a baby boom after World War II, but only some of them also underwent a substantial mobilization of female labor during the war. Our theory predicts that countries with a larger wartime increase in the female labor force should also experience larger baby booms. We assess this prediction by comparing the baby boom in two sets of countries: the Allied countries that, like the United States, did not fight on their own soil (Australia, Canada, and New Zealand), and the major European countries that remained neutral in the war (Ireland, Portugal, Spain, Sweden, and Switzerland).[33] The results confirm our hypothesis. The Allied countries mobilized a substantial fraction of working-age men for the war, which resulted in a large increase in female labor-force participation. Subsequently, all of the Allied countries experienced a baby boom and baby bust that is remarkably similar to that of the United States. In contrast, in the neutral countries the war did not mark a watershed for female labor-force participation, and the post-war baby boom was of a much smaller magnitude than in the Allied countries. In what follows, we present more detailed information on the involvement of these two groups of countries in the war and their subsequent fertility experience.

Australia joined the war on September 3, 1939, the same day Britain and France declared war on Germany. In September 1939, only 14,903 men were enlisted in the Royal Australian Navy, the Australian Military Forces, and the Royal Australian Air Force. Enlistment grew rapidly, however: by November 1941, 364,874 men were enlisted, and within less than a year the size of the armed forces nearly doubled to 634,645 in August 1942. During the years 1942–1945, between 23 and 27 percent of all males age 15–64 were serving in the armed forces.[34] New Zealand joined the war on the same day Australia did. In September 1939, 20,806 men were serving, but this number grew rapidly to a peak of 154,549 in July

[33] Australia was subject to some aerial bombing and naval shelling, but destruction was on a much smaller scale than in the United Kingdom.

[34] Authors' calculation using series WR24, POP211 and POP274 in Vampley (1987).

Illustration 2: Australian Government World War II Recruitment Poster

1942. During the years 1942–1944, between 30 and 37 percent of all males age 20–59 were serving in the armed forces.[35] Canada joined the war seven days after Australia and New Zealand. At that time, only 9,000 individuals were in the armed services. By 1941 enlistment had reached 296,000, and the peak was reached in 1944 with 779,000 men under arms. At this time, nearly 19 percent of all males age 15-64 were serving in the armed forces.[36]

As in the United States, the mobilization of men led to a large increase in female employment during the war, with active encouragement by government campaigns (see Illustration 2). For the generation of women old enough to work during the war, the increase in labor-force participation persisted in the following decades. For example, in Canada the labor-force participation of women aged 35–64 increased by more than 30 percent between 1941 and 1951 and by another 50 percent between 1951 and 1961. In contrast, the participation rate of women aged 25–34 in 1951 was down nearly 10 percent compared to 1941, and by 1961 it exceeded the 1941 level by merely 5 percent.[37] This pattern closely resembles our

[35] Authors' calculation using Tables II.4 and VIII.17 in Bloomfield (1984).

[36] Authors' calculation using series A32-A41 and C48 in Urquhart and Buckley (1965).

[37] Source: Historical Estimates of the Canadian Labour Force, 1961 Census Monograph, Statistics Canada, Catalogue 99-549. The data for New Zealand and Australia are less detailed. For

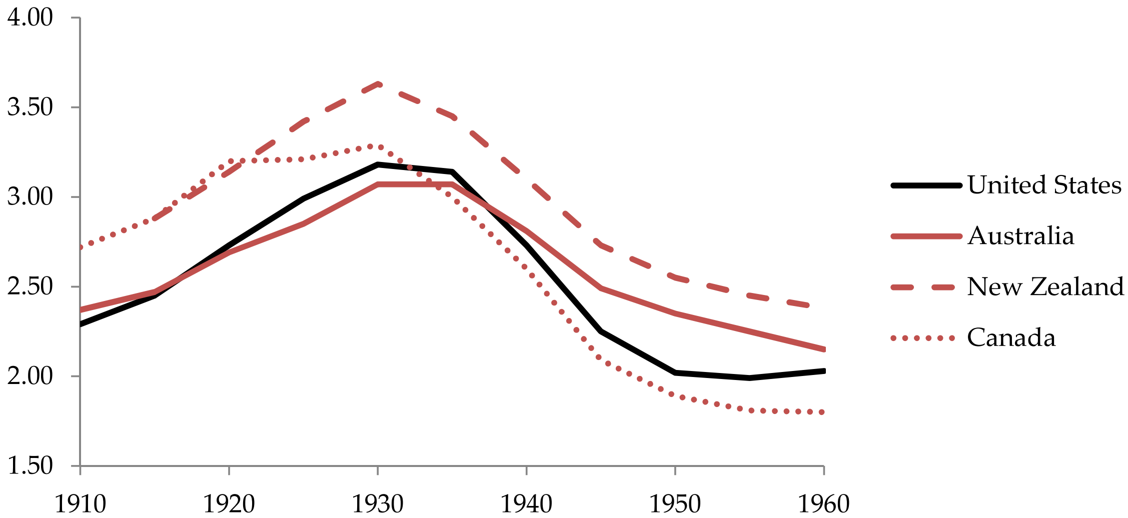

Figure 13: Completed Fertility Rate by Birth Cohort in Australia, Canada, New Zealand, and the United States

findings for the United States.

The fertility dynamics in the Allied countries in the post-war period display a striking resemblance to the United States. Figure 13 displays the completed fertility rate in the United States, Canada, Australia, and New Zealand for women born between 1910 and 1960.[38] In all four countries, the completed fertility rate increased steadily from the cohorts born in the 1910s to those born in the early 1930s. Subsequently, completed fertility declined in all four countries. In the United States and Australia, the completed fertility rate peaks for women born in 1932. In Canada the peak is reached with the 1931 birth cohort, and in New Zealand with the 1930 cohort. The similarity of the fertility experience of these countries concerns not only the timing but also the magnitude of the baby boom. Measured as the absolute difference between the completed fertility rate of women born in 1913 and women born in the early 1930s, the size of the baby boom equals 0.8 in the United States and New Zealand, 0.79 in Australia, and 0.48 in Canada.

Australia, Beaton (1982) reports that the total number of employed women during the war rose by nearly 200,000 between 1939 and 1943, and while it had dropped by nearly 70,000 from that peak by 1946, by 1948 the total number of employed women had risen above the 1943 peak by 4000.

[38]Data on completed fertility rates were kindly provided by Jean-Paul Sardon of the Observatoire Démographique Européen, which maintains a database on fertility in Europe. Some of these

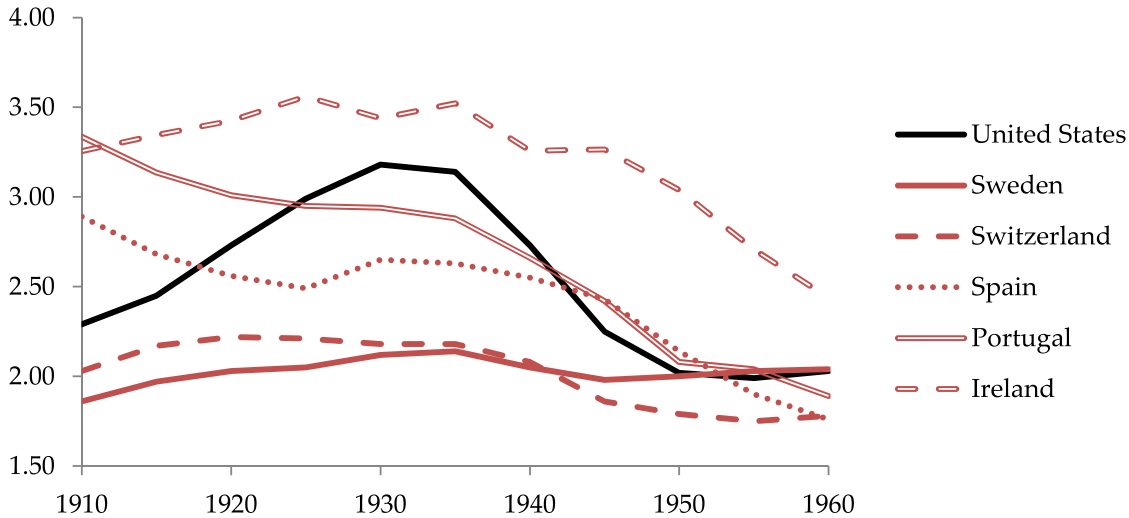

Figure 14: Completed Fertility Rate by Birth Cohort in Ireland, Portugal, Spain, Sweden, Switzerland, and the United States

Five major European countries remained officially neutral in World War II: Ireland, Portugal, Spain, Sweden, and Switzerland. While there was some wartime mobilization even in these countries (in particular in Switzerland), these countries did not experience a substantial increase in female employment during the war.[39] Our mechanism for the baby boom therefore does not apply to these countries, and consequently we would expect to observe smaller post-war baby booms (which must then be due to other mechanisms). Figure 14 shows the completed fertility rate in the five neutral countries in comparison to the United States. The figure shows that Portugal did not experience any baby boom at all. In Ireland there is a small rise in fertility between the 1910 and 1925 cohorts, but both the initial rise and the subsequent decline are more gradual than in the United States, without a sharp boom-bust pattern. Fertility also went up in Spain, Sweden, and Switzerland, but again much less so than in the United States. Among the neutral countries, Ireland experienced the largest increase in fertility, but even here the size of the baby boom (in terms of the increase in the completed fertility rate) is only 0.3, less than half of the increase in the United

data are published in Sardon (1990) and Sardon (2006).

[39]Nor did female participation rise quickly after the war. For example, Sweden and Switzerland do not show any marked increase in the female labor-force participation rate between 1940 and 1960.

States.

We focus on neutral countries as the control group because the other major industrialized countries at the time (Italy, Japan, France, Germany, and the United Kingdom) all experienced massive wartime destruction as well as loss of life, which in itself is likely to have had a major impact on subsequent fertility. For what it's worth, post-war fertility in these countries does not display the pronounced baby boom and baby bust that we observe in the United States and the other Allies. In Japan, fertility recovered right after the war, but then dropped sharply from the early 1950s. The European countries experienced baby booms that were substantially smaller and occurred substantially later than in the United States, with peak total fertility rates between 2.6 and 2.9 in 1964 (in the United States, the peak of the baby boom was in 1957 with a fertility rate of 3.8). In terms of female labor force participation, only the United Kingdom saw widespread female employment during the war and rising overall female participation afterwards.[40] We conjecture that the much larger loss of capital and life explains the difference in post-war fertility regimes between the United Kingdom and the other Anglo-Saxon Allies.

In sum, the international evidence suggests that our mechanism is quantitatively important for explaining the baby boom and baby bust of the 1950s and the 1960s. Allied countries that were shielded from wartime destruction and that experienced a large increase in female employment during the war also had much larger subsequent baby booms than neutral countries. In our comparison, we have have focused on the completed fertility rate as a measure of fertility, because it corresponds most closely to the predictions of our model. However, carrying out the comparison in terms of the total fertility rate would lead to the same conclusions.

6 Relationship to Literature

The perhaps most widely known explanation for the baby boom is Easterlin's (1961) relative income hypothesis.[41] Easterlin postulates that fertility decisions

[40]The Axis countries promoted traditional role models and relied much less on female labor.

[41]Also well known is what might be termed the "catch-up fertility" hypothesis, i.e., the idea that fertility rates rose after the war because couples were making up for babies they were not

are driven by the gap between couples' actual and expected material well-being. Applying this theory to the U.S. baby boom, Easterlin argues that people who grew up during the Great Depression had low material aspirations. Overwhelmed by the prosperity of the post-war years, they increased their demand for children. One of the problems with this explanation is that the timing is not quite right.[42] As we documented above, most of the baby boom was accounted for by young mothers aged 20–24. During the baby boom fertility peaked in 1957. Mothers who were 20–24 years old in 1957 were born between 1933 and 1937, and spent much of their childhood during the prosperous post-war period.[43]

Greenwood, Seshadri, and Vandenbroucke (2005) propose an alternative theory based on improvements in household technology. They argue that the widespread diffusion of appliances such as refrigerators, washers, dishwashers, and electric stoves enabled women to run their households in much less time than before, which lowered the time cost of raising children.[44] This theory is complementary to ours in the sense that in each case the focus is on the opportunity cost of having children, albeit from different ends: In Greenwood et al. the direct cost of having a child declines, while in our model it is the opportunity cost of time (i.e., young women's wage) that goes down. We believe both these aspects to be relevant. One observation that supports the relative importance of our theory is that most of the baby boom is accounted for by young women. If the baby boom was exclusively due to a general lowering of the cost of having children, we would expect to see a substantial increase in fertility at all ages. The data suggest that

able to have during the war. However, while this mechanism probably contributed to the spike in fertility in 1946 and 1947, the literature has long recognized that it cannot explain the main phase of the baby boom in the 1950s. Most of the baby-boom mothers were too young to be married during the war. More importantly, the data on completed fertility show that women increased their lifetime fertility during the period, which would not be the case if the baby boom solely represented a shift in the timing of births from during to after the war (see Figure 2).

[42] This was pointed out by Greenwood, Seshadri, and Vandenbroucke (2005).

[43] A related possibility is that the state of the economy has an immediate impact on fertility, rather than working with a lag of one generation as in Easterlin's hypothesis. Along these lines, Jones and Schoonbroodt (2011) argue that economic shocks such as the Great Depression and World War II can have large affects on fertility behavior, potentially explaining a significant fraction of the trough in fertility in the 1930s and the subsequent recovery during the baby boom.

[44] Related papers that also attribute part of the baby boom to a decline in the cost of children are Murphy, Simon, and Tamura (2008), who argue that suburbanization lowered the cost of space, and Albanesi and Olivetti (2010), who focus on medical progress that lowered the risk of childbirth.

there was a decline in a component of the opportunity cost that applies only to young mothers. This is exactly what happens in our model: the key margin is the opportunity cost of time at the transition from working to motherhood, which is operative only at the beginning of the childbearing period. A second advantage of our theory is that it does better at accounting for the baby bust, i.e., the sharp decline of fertility in the 1960s.

Our theory shares with Butz and Ward (1979) the emphasis on relative female wages as a key determinant of the opportunity cost of having children. An important distinction is that in our theory the earnings potentials of young and old women have opposite implications for fertility choices, whereas Butz and Ward do not make this distinction.[45] Moreover, we provide a general-equilibrium model in which relative male and female wages are endogenously determined, which is essential for our main argument, namely the link between the demand for female labor during World War II and the subsequent baby boom.

Zhao (2011) argues that the increase in income taxes after the war to pay down wartime debt was an important cause of the baby boom. In particular, higher taxes lowered the opportunity cost of child-rearing and therefore increased fertility. While fiscal changes play a role in our theory as well, the proportional changes in taxes that Zhao focuses on do not affect fertility in our model. Moreover, the fiscal channel alone cannot explain the baby bust, given that taxes did not come down after the peak of the baby boom.

The structure of our model of fertility choice is related to Galor and Weil (1996), who also provide a general-equilibrium model where couples jointly decide on fertility and labor supply and where the opportunity cost of children is determined by the relative female wage. However, Galor and Weil focus on the long-run trend of declining fertility and do not discuss the baby boom. In addition, unlike Galor and Weil we develop a life-cycle model where the interaction between successive cohorts is key for the economic mechanism. Our emphasis on the timing of fertility is shared with Caucutt, Guner, and Knowles (2002), who use an integrated model of the marriage market, female labor supply, and fertil-

[45] One indication that this matters is the finding in Macunovich (1995) that the original Butz and Ward results cannot be replicated with more recent updated data.

ity to explain patterns of fertility timing in the United States. Life-cycle models of fertility and female labor-force participation have also been developed and estimated in the labor literature (see for example Moffit 1984 and Eckstein and Wolpin 1989). One feature that is important both in this literature and our theory is the endogenous accumulation of work experience. However, the papers in the labor literature focus on the estimation of partial-equilibrium choice models, and therefore lack the general equilibrium aspect that is essential for our mechanism.

We also build on the literature that analyzes the role of World War II in explaining the rise in female employment. Goldin (1991) reports that 25 percent of the working women in the age group 27–51 in 1951 were women who did not work in December 1941 but worked in March 1944 and January 1951, and are therefore likely to have entered the labor force due to the war. Among the women who did not work in December 1941 but worked in March 1944, 65 percent were still working in 1951. Given that our quantitative model provides a good match for female labor-force participation during and after the war, our analysis is consistent with these findings.[46]

An important question is whether the war may have had an impact on female employment that went beyond the impact on individual employment histories that Goldin (1991) focuses on. An argument of this kind is presented by Fernández, Fogli, and Olivetti (2004), who argue that one factor that held back female employment is husbands' prejudice against working wives. The extent of this prejudice, in turn, depends on whether a husband's own mother was working. Given this mechanism, the demand for female labor during the war increased married women's labor-force participation one generation later, when the sons of the working mothers of the war got married. More generally, it has been argued that simply observing more married women work will reduce prejudice against and misinformation about working women. Along these lines, Hazan and Maoz (2002), Fernández (2012), and Fogli and Veldkamp (2011) have developed models that give rise to the S-shape dynamics in female labor-force participation that characterize the data. Accounting for mechanisms of this kind would lead to an even higher estimate of the impact of World War II on female employment.

[46] See also Clark and Summers (1982) for additional evidence supporting an important role of World War II for the rise in female employment.

7 Conclusions

In this paper, we have proposed a simple theory to argue that the shock of World War II can account for a substantial part of the rise and fall of U.S. fertility through the post-war baby boom and baby bust. Earlier research has dismissed a causal link between the war and increased fertility, mainly because the baby boom extended for 15 years after the war and is too large to be explained solely by "catch up" fertility. Our theory, however, does not rely on "catch up" fertility, but on the implications of the war for the female labor market.

We show that if female labor supply is persistent, a one-time demand shock for female labor leads to long lasting, asymmetric effects on the labor supply of younger and older women. World War II was a huge demand shock for female labor. As a consequence, the war generation of women continued to work throughout the baby boom period, whereas younger women were crowded out from the labor market and had more children instead. The labor market channel is further amplified by the fiscal consequences of the war, and in particular the persistent rise in labor taxation. Our quantitative analysis suggests that these mechanisms can account for a major portion of the rise and fall in completed fertility rates during the baby boom and baby bust periods.

One aspect of the data that our theory does not account for is the substantial difference between the increase in total fertility rates versus completed fertility rates during the baby boom (see Figures 1 and 2). This discrepancy is due to the timing of births: at the height of the baby boom older cohorts of women were having children late, while younger women were having them early. This observation suggests that there was some factor present that induced women to have their babies *at the same time* as other women, leading to a coordination in fertility that increased period fertility rates in the late 1950s well above the realized fertility rate of any given cohort. A related observation is that even though much of the baby boom is due to the young women who we concentrate on here, during the 1950s we observe at least some increase in fertility for almost any group of women, irrespective of age, labor force status, or place of residence. To us, these observations suggest the presence of social externalities in child rearing that made the period fertility rate highly elastic.[47] For example, when many families in a community are already having children, it may be easier to arrange informal child care or children's activities, which lowers the cost of having another child. Modeling such externalities and using them to explain the full increase in period fertility during the baby boom is an important challenge for future research.

References

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

Albanesi, Stefania, and Claudia Olivetti. 2010. "Maternal Health and the Baby Boom." Unpublished Manuscript, Columbia University.

Beaton, Lynn. 1982. "The Importance of Women's Paid Labour: Women at Work in World War II." In *Worth Her Salt: Women at Work in Australia*, edited by Margaret Bevege, Margaret James, and Carmel Shute, 87–91. Sydney: Hale and Iremonger.

Bloomfield, Gerald T. 1984. *New Zealand: A Handbook of Historical Statistics*. Boston: G.K. Hall.

Butz, William P., and Michael P. Ward. 1979. "The Emergence of Countercyclical U.S. Fertility." *American Economic Review* 69 (3): 318–28.

Caucutt, Elizabeth M., Nezih Guner, and John Knowles. 2002. "Why Do Women Wait? Matching, Wage Inequality, and the Incentives for Fertility Delay." *Review of Economic Dynamics* 5 (4): 815–855.

Chesnais, Jean Claude. 1992. *The Demographic Transition*. Oxford: Oxford University Press.

Clark, Kim B., and Lawrence H. Summers. 1982. "Labour Force Participation: Timing and Persistence." *Review of Economic Studies* 49 (5): 825–44.

Cooley, Thomas F., and Edward C. Prescott. 1995. "Economic Growth and Business Cycles." Chapter 1 of *Frontiers of Business Cycle Research*, edited by Thomas F. Cooley. Princeton: Princeton University Press.

Doepke, Matthias. 2004. "Accounting for Fertility Decline During the Transition to Growth." *Journal of Economic Growth* 9 (3): 347–83.

Easterlin, Richard A. 1961. "The American Baby Boom in Historical Perspective." *American Economic Review* 51 (5): 869–911.

[47] Such externalities have been documented for the labor force participation of mothers, see Maurin and Moschion (2009).

Eckstein, Zvi, and Kenneth I. Wolpin. 1989. "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience." *Review of Economic Studies* 56 (3): 375–90.

Fernández, Raquel. 2012. "Culture as Learning: The Evolution of Female Labor Force Participation over a Century." Forthcoming, American Economic Review.

Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti. 2004. "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." *Quarterly Journal of Economics* 119 (4): 1249–99.

Fogli, Alessandra, and Laura Veldkamp. 2011. "Nature or Nurture? Learning and the Geography of Female Labor Force Participation." *Econometrica* 79 (4): 1103–38.

Galor, Oded, and David N. Weil. 1996. "The Gender Gap, Fertility, and Growth." *American Economic Review* 86 (3): 374–87.

Goldin, Claudia. 1990. *Understanding the Gender Gap: An Economic History of American Women*. Oxford University Press.

———. 1991. "The Role of World War II in the Rise of Women's Employment." *American Economic Review* 81 (4): 741–56.

Greenwood, Jeremy, Ananth Seshadri, and Guillaume Vandenbroucke. 2005. "The Baby Boom and Baby Bust." *American Economic Review* 95 (1): 183–207.

Haveman, Robert, and Barbara Wolfe. 1995. "The Determinants of Children's Attainments: A Review of Methods and Findings." *Journal of Economic Literature* 33 (4): 1829–78.

Hazan, Moshe, and Yishay D. Maoz. 2002. "Women's Labor Force Participation and the Dynamics of Tradition." *Economics Letters* 75 (2): 193–98.

Heston, Alan, Robert Summers, and Bettina Aten. 2006. "Penn World Table Version 6.2." Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hill, Russell C., and Frank P. Stafford. 1980. "Parental Care of Children: Time Diary Estimates of Quantity Predictability and Variety." *Journal of Human Resources* 15 (2): 219–39.

Hobbs, Frank, and Nicole Stoops. 2002. *Demographic Trends in the 20th Century*. Census 2000 Special Reports. Washington: U.S. Census Bureau.

Joines, Douglas H. 1981. "Estimates of Effective Marginal Tax Rates on Factor Incomes." *Journal of Business* 54 (2): 191–226.

Jones, Larry E., and Alice Schoonbroodt. 2011. "Baby Busts and Baby Booms: The Fertility Response to Shocks in Dynastic Models." Unpublished Manuscript, University of Minnesota.

Jones, Larry E., and Michèle Tertilt. 2008. "An Economic History of Fertility in the United States: 1826–1960." Chapter 5 of *Frontiers of Family Economics*, edited by Peter Rupert. Bradford: Emerald.

Lagerlöf, Nils-Petter. 2006. "The Galor-Weil Model Revisited: A Quantitative Exercise." *Review of Economic Dynamics* 9 (1): 116–42.

Macunovich, Diane J. 1995. "The Butz-Ward Fertility Model in the Light of More Recent Data." *Journal of Human Resources* 30 (2): 229–55.

Maurin, Eric, and Julie Moschion. 2009. "The Social Multiplier and Labor Market Participation of Mothers." *Americal Economic Journal: Applied Economics* 1 (1): 251–72.

McGrattan, Ellen, and Lee E. Ohanian. 2010. "Does Neoclassical Theory Account for the Effects of Big Fiscal Shocks? Evidence from World War II." *International Economic Review* 51 (2): 509–32.

Moffit, Robert. 1984. "Profiles of Fertility, Labour Supply and Wages of Married Women: A Complete Life-Cycle Model." *Review of Economic Studies* 51 (2): 263–78.

Mulligan, Casey B. 1998. "Pecuniary Incentives to Work in the United States During World War II." *Journal of Political Economy* 106 (5): 1033–77.

Murphy, Kevin M., Curtis Simon, and Robert Tamura. 2008. "Fertility Decline, Baby Boom, and Economic Growth." *Journal of Human Capital* 2 (3): 262–302.

Ohanian, Lee E. 1997. "The Macroeconomic Effects of War Finance in the United States: World War II and the Korean War." *American Economic Review* 87 (1): 23–40.

Olivetti, Claudia. 2006. "Changes in Women's Hours of Market Work: The Role of Returns to Experience." *Review of Economic Dynamics* 9 (4): 557–87.

Piketty, Thomas, and Emmanuel Saez. 2007. "Income Inequality in the United States, 1913-2002." Chapter 5 of *Top Incomes over the Twentieth Century: A Contrast Between European and English Speaking Countries*, edited by A. B. Atkinson and T. Piketty. Oxford: Oxford University Press.

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

Sardon, Jean-Paul. 1990. *Cohort Fertility in Member States of the Council of Europe*. Population Studies, No. 21. Strasbourg: Council of Europe.

———. 2006. "Fertility in the Developed English-speaking Countries Outside Europe: Canada, United States, Australia and New Zealand." *Population* 61 (3): 267–92.

Siu, Henry E. 2008. "The Fiscal Role of Conscription in the U.S. World War II Effort." *Journal of Monetary Economics* 55 (6): 1094–1112.

Urquhart, Malcolm C., and Kenneth A.H. Buckley. 1965. *Historical Statistics of Canada*. Cambridge: Cambridge University Press.

U.S. Department of Commerce, Bureau of the Census. 1975. *Historical Statistics of the United States, Colonial Times to 1970*. Bicentennial Edition. Washington: U.S. Department of Commerce, Bureau of the Census.

Vampley, Wray, ed. 1987. *Australians: Historical Statistics*. Broadway, New South Wales: Fairfax, Syme and Weldon.

Whelpton, Pascal K. 1964. "Trends and Differentials in the Spacing of Births." *Demography* 1 (1): 83–93.

Zhao, Kai. 2011. "War Debt and the Baby Boom." Unpublished Manuscript, University of Western Ontario.

A Appendix

A.1 Probit Regression Results

Table 5 displays regression results analogous to those in Tables 2 and 3 using ordered probit (Children under Age 5) and probit (Employed and Ever Married) regression.

A.2 Fertility Data

Data on total fertility rates (TFR) in Figures 1 and 7 are taken from Chesnais (1992), Tables 2A.3 and 2A.4, pp. 545–548. Data on completed fertility rates in Figures 2 and 6 were kindly provided by Jean-Paul Sardon of the Observatoire Démographique Européen, which maintains a database on fertility in Europe. Some of these data are published in Sardon (1990) and Sardon (2006) (See also Jones and Tertilt (2008), Table A1, p. 56). The average age at first birth is computed from data on first birth rate by age, taken from the Historical Statistics of the United States, Millennial Edition, Vol. 1, Table Ab150–215, pp. 412–413.

Table 5: Impact of WWII Mobilization Rates on Fertility, Female Labor Supply, and Probability of Marriage (Coefficient Estimates from Probit and Ordered Probit Regressions for Variable "Mobilization Rate $\times$ 1960")

Dependent variable	Regression				
	(1)	(2)	(3)	(4)	(5)
	Age 25-35 (N = 243554)				
Children under Age 5	1.760	0.971	0.746	3.092	2.840
	(0.307)	(0.266)	(0.253)	(0.519)	(0.564)
Employed	-2.371	-1.086	-0.419	-2.360	-1.293
	(0.423)	(0.379)	(0.360)	(0.774)	(0.773)
Ever Married	0.628	0.674		2.093	
	(0.414)	(0.425)		(0.636)	
	Age 45-55 (N = 191715)				
Employed	-0.366	0.423	0.556	1.135	1.324
	(0.258)	(0.281)	(0.271)	(0.926)	(0.846)
Education and Farm Controls	no	yes	yes	yes	yes
Marital Status Controls	no	no	yes	no	yes

Notes: Standard errors (in parentheses) are adjusted for clusters of state of residence and year of observation. Estimates are from separate regressions of pooled micro data from 1940 and 1960 census. Regressions 1-3 are probit and ordered probit and 4-5 use instruments. Each outcome variable is regressed on the WWII mobilization rate interacted with a 1960 year indicator variable, indicator variables of observation year, age, race, state of residence, state/country of birth. All indicator variables except state/country of birth and state of residence are also interacted with the 1960 year indicator variable. Instrumental variables used in regressions 4-5 are: 1940 male share ages 13-24 interacted with a 1960 year indicator variable, 1940 male share ages 25-34 interacted with a 1960 year indicator variable, and 1940 male share German interacted with a 1960 year indicator variable. All data is weighted using census person weights.

A.3 Labor Supply and Wages

Statistics on labor supply and wages are computed from census data. Specifically, we use data from the 1 percent Integrated Public Use Microsample (IPUMS) of the Decennial Census for the decades 1940 to 1990. For 1940, 1950, 1960, 1980 and 1990, we use the general 1 percent sample. For 1970 we use the Form 2 Metro sample. The data are weighted using the appropriate weighting scheme (see Ruggles et al. 2010). We restrict our attention to individuals aged 20–60, living in non-farm households, and whose group quarter status is equal to 1, "Households under the 1970 definition."

Total hours worked in the previous year is computed by multiplying weeks worked last year (WKSWORK1) by hours worked last week (HRSWORK1). In 1960 and 1970, Census information on weeks worked last year and hours worked last week are reported only in intervals (WKSWORK2 and HRSWORK2, respectively). Therefore, for these decades, weeks worked last year and hours worked last week are assigned the midpoint value of each interval as in Fernández, Fogli, and Olivetti (2004).

The variable weeks worked last year (WKSWORK1) is not comparable across all years. Specifically, as noted by Ruggles et al. (2010), in 1940 "It was up to respondents to determine precisely what "full-time" meant in their specific locality, occupation, and industry. If respondents did not know how many hours should be regarded as a full-time week, enumerators were instructed to suggest that 40 hours was a reasonable figure. In essence, respondents were to estimate how many hours they had averaged per week, multiply this figure by 52 weeks, then divide by about 40."

To assure comparability between 1940 and subsequent Census years, we took the following steps. For individuals who reported 52 weeks in the previous year or less than 52 weeks in the previous year but 40 or more hours in the previous week, we left the annual hours unchanged (i.e., WKSWORK1 times HRSWORK1). For those who reported less than 52 weeks in the previous year and less than 40 hours in the previous week, we computed annual hours as weeks worked last year times 40 (i.e., WKSWORK1 times 40).

The measure of labor supply reported in the paper is the ratio between the mean annual hours worked by women to the mean annual hours worked by men in the same group. This measure can be interpreted as a full-time equivalent labor-force participation rate. When comparing model to data (see Figure 9) we use data for all women for the ages 20–32 but for married women for the ages 33-60, because our model does not allow for the possibility of spinsterhood.

For wages and the gender gap, we use the information on wage and salary income (INCWAGE). N/A code (999999) is treated as a missing value. Following Acemoglu, Autor, and Lyle (2004), top-coded values are imputed as 1.5 times the censored value. To obtain hourly wages, INCWAGE is divided by the total hours worked in the previous year. The relative wage of women to men, i.e., 1 minus the gender gap is computed as the ratio of the mean wage of women to the mean wage of men in the same group. For the overall gender gap, we use average wages for ages 20–60. For the gender gap for young women (i.e., before child bearing) we use data on single women aged 20–24 (see Figure 10). While formally in our model all women start out already married, the pre-child-bearing period is best interpreted as corresponding to single life in the data. Empirically, marriage and having one's child were closely related events during the period. In addition, using data for single women is less subject to selection problems, as the vast majority of young single women were working.

A.4 The Accumulation of Work Experience

To calibrate the experience accumulation function, we estimate an earnings equation using data from the 1940 census. The estimation equation is:

$$\ln w_i = \alpha + \omega_0 \text{ education}_i + \omega_1 \text{ experience}_i + \omega_2 \left(\text{ experience}_i \right)^2 + \epsilon_i. \tag{1}$$

The equation is estimated for men aged 20–60 using a Heckman selection model. We assume that the selection into the labor force depends on education, marital status, and the number of children under the age of 5. Given that actual work experience is not available, we follow the standard in the labor literature and compute experience as:

$$\text{experience} = \text{age} - \text{education} - 6.$$

We obtain estimates of the return-to-experience parameters of $\omega_1 = 0.05$ and $\omega_2 = -0.00053$. Given that in the model people start work at age $j = 1$ and that a model period corresponds to 2.5 years, we choose the return to experience such that the efficiency units of labor supplied by a man of age j are given by:

$$e^m_{t,j} = \exp(\omega_1 \cdot 2.5(j-1) + \omega_2 \cdot (2.5(j-1))^2).$$

Here we normalize $e^m_{t,1} = 1$. Iterating this expression to age $j+1$ and rearranging yields:

$$e^m_{t,j+1} = \exp(2.5\omega_1 + \omega_2(12.5j - 6.25))e^m_{t,j},$$

so that:

$$\eta_{m,j} = \exp(2.5\omega_1 + \omega_2(12.5j - 6.25)) - 1.$$

Substituting the estimates for ω_1 and ω_2 gives:

$$\eta_{m,j} = \exp(0.125 - 0.00053(12.5j - 6.25)) - 1.$$

We also assume that the return to experience for women and men is the same, $\eta_{f,j} = \eta_{m,j}$ for all j. We then choose the return to age ν such that in the pre-war balanced growth path, at age 32.5 (when women reach the end of the fecund period) the productivity of working women is larger by a factor of 1.42 than at age 20. This factor is obtained by estimating an earnings function for women (using the same functional form as used for men in (1)) and predicting women's wages at ages 20 and 32.5. The procedure yields a return to age (per model period) of $\nu = 0.003$ (thus, the return to age is close to zero).

A.5 The Child Care Cost Function

The level parameter ϕ of the child-care cost function is pinned down using data on the total private cost of children from Haveman and Wolfe (1995), as described in the main text. However, Haveman and Wolfe (1995) do not report information which can be used to back-up the curvature parameter ψ. We therefore use time use data to set ψ. We estimate ψ by running the regression $\ln y_i = \omega_0 + \psi \ln n_i + \epsilon_i$ on time use data. The data come from the American Heritage Time Use Study. Specifically, we follow Hill and Stafford (1980) and use the 1975–1976 American's Use of Time survey. This is a panel study designed and administered by the Survey Research Center at the University of Michigan.[48] We also followed Hill and Stafford (1980) in defining child care as the sum of minutes care for infant, minutes care for older child, minutes medical care for child, minutes play with child, minutes supervise homework, minutes read to/talk to child, and minutes other child care. Restricting attention to women of all marital statuses who live in urban areas, we obtain $\psi = 0.3024$. Similarly, restricting attention to married women, we obtain $\psi = 0.3509$. We use the average of these estimates and fix the curvature parameter at a value of 0.33.

In addition to the two parameters ϕ and ψ, we also need to fix the additional time cost associated with a birth, κ. In the time use data described above, we find that mothers with one child in the age group 0–3 spent somewhat more than twice as many minutes per day than mothers with one child who is older than 3 years old. Since time costs make up only a fraction of the total private cost of children, we set $\kappa = 0.5\phi$.

A.6 U.S. Mobilization for World War II

Data on the mobilization of American men to World War II are taken from U.S. Department of Commerce, Bureau of the Census (1975), series Y904, "Military Personnel on Active Duty." To conversion of the absolute numbers to rates we divide this series by the male population in the age group 20–59. These numbers come from Hobbs and Stoops (2002), Table 5: "Population by Age and Sex for the United States: 1900 to 2000 Part A." Since the population in this age group is available only on a decennial basis, we assume a constant growth rate between 1940 and 1950. This procedure yields mobilization rates of 0.013, 0.049, 0.104, 0.242, 0.303, 0.318, 0.079 and 0.041 for the years 1940–1947, respectively. Hence, a reduction of 30 percent of men availability for one period is based on the average mobilization rate during the 1943-1945 period. Note that this is a conservative reduction as we disregard the decline in men's availability during the 1941–1942 period.

[48] The data are available online at: http://www.timeuse.org/ahtus and were downloaded from this web-site on September 20th, 2007. The 1975–1976 survey was designed as a nationally representative sample of households and sampled both respondents, and, if the respondent was in a couple, the spouse or partner. Four waves of the survey were carried out to represent all seasons of the year and all days of the week. The study collected most information from one person per household. However, if the diarist had a spouse, the spouse was asked to complete a cut-down version of the diary and questionnaire.