Marriage and Misallocation: Evidence from 70 Years of U.S. History*

Jay Euijung Lee†

Stockholm University and Centre for Economic Performance

January 5, 2024

Abstract

The traditional expectation that married women should be homemakers precludes some married women from pursuing their comparative advantage in the labor market and discourages education. I quantify the aggregate economic consequences of gender norms in marriage, accounting for selection into education, marriage, and labor force participation. Relative to similar single women, married women of 1940 faced a norms wedge that acted as a 44% tax on market wage. By 2010, the norms wedge fell to 14%. Had norms remained at the level of 1940, the total market and home output of the current day would have been 5.5% lower.

Keywords: Marriage, gender norms, misallocation, female labor force participation

*I wish to especially thank Maitreesh Ghatak, Guy Michaels, Oriana Bandiera, Robin Burgess, Gharad Bryan, Tim Besley for their guidance. I also thank Alexia Delfino, Alessandro Gavazza, Ethan Ilzetski, Camille Landais, Stephen Machin, Alan Manning, Isabelle Mejean, Ameet Morjaria, Rachel Ngai, Barbara Petrongolo, Steve Pischke, John Van Reenen, Silvana Tenreyro, and Horng Chern Wong for valuable comments and discussions. This research is supported by a grant of the French National Research Agency (ANR), "Investissements d'Avenir" (LabEx Ecodec/ANR-11-LABX-0047). Some of the data used in this analysis are derived from Sensitive Data Files of the GSS, obtained under special contractual arrangements designed to protect the anonymity of respondents. These data are **not** available from the author. Persons interested in obtaining GSS Sensitive Data Files should contact the GSS at **GSS@NORC.org**.

†email: Jay-Euijung.Lee@su.se

1 Introduction

Many women shift their time from the labor market to home production upon marriage (Lundberg and Pollak, 2007). If there are efficiency reasons to specialize between income-generating activity and home production, this shift might enhance productivity (Becker, 1981, 1991; Pollak, 2013). On the contrary, the shift might also represent misallocation. The need to fulfill the traditional gender role of the homemaker, who stays home to look after the household, might prevent some married women from following their comparative advantage and working in the labor market.

By how much do traditional gender norms in marriage constrain aggregate output? This paper quantifies the effect on aggregate output of the change in the "homemaker" gender role, in the U.S. between 1940 and 2010. While there is ample micro evidence on how gender roles curtail the market work of women (e.g. Field et al., 2020; Couprie et al., 2020; Bedi et al., 2021), little is known about the aggregate implications of gender norms. Existing papers on the aggregate output implications of gender differences such as labor market and educational market discrimination (Hsieh et al., 2019) and nonmarket time (Erosa et al., 2022) do not distinguish between *married and single* women, despite crucially disparate labor market outcomes and trends over time (Goldin, 2006; Blau and Kahn, 2007). The literature on the intersection of Macroeconomics and Family Economics does place emphasis on how marriage and family are critical for understanding various aggregate outcomes – business cycles, income inequality, fertility, labor supply, savings, long-run development, and political institutions (see Doepke and Tertilt (2016) for an overview) – yet few build the connection between gender roles and output. My contribution is to fill this gap.

Two main challenges arise in quantifying the output effects of gender norms. The first is on the measurement of norms. Measures of norms or attitudes are typically derived from attitudinal or opinion surveys or more ingenious sources such as folklore (Goussé, Jacquemet, and Robin, 2017; Bursztyn, González, and Yanagizawa-Drott, 2020; Michalopoulos and Xue, 2021), but the way to connect them to output in a general equilibrium framework remains obscure. To do so, norms would rather need to be measured either from within a model or have direct linkages with model ingredients. The second challenge is that such a model would need to treat carefully how key determinants of output, such as human capital accumulation, interact with norms. Whether more conservative norms against married women working would lower labor market returns to education for women is far from obvious if fewer women choose to get married.

I develop a model featuring education, marriage, and labor supply choices to quantify the consequences of gender roles in marriage. Gender roles are measured, through the lens of the model, as a composite force that makes the labor force participation of married women diverge from that of similar single women,[1] besides the wage differentials.[2] The model is matched to the education, marriage, and labor force participation patterns in the U.S. decennial census, decade by decade, to track how the magnitude of gender roles changes over time. The model-derived measures of norms are highly correlated with attitudinal survey responses, and the model structure is validated externally by a reduced-form analysis.

My theoretical contribution is to build a tractable general equilibrium model with closed-form solutions, while maintaining flexibility to account for key secular trends in the gender wage gap, gender differences in home productivity, propensity of marriage, assortativeness of marriage matching by education, and educational attainment. The closed-form solutions are powerful as they lay out precisely how the three decisions are interconnected and elucidate the parameter identification procedure. To do so, I adopt a tractable form of selection into labor activity by individuals with heterogeneous abilities, commonly used in the Trade literature (Eaton and Kortum, 2004; Hsieh et al., 2019). The novelty of the model is to extend this form of selection, previously only used to study *individual* choices, into a model of *household* decision-making, and to seamlessly integrate the resulting household economic utilities into standard models of educational choice and marriage market matching (Choo and Siow, 2006; Chiappori, Salanié, and Weiss, 2017; Chiappori, Costa-Dias, and Meghir, 2018).

Individuals in my model make three sets of choices over the course of their life cycle. First, individuals choose their level of education as a forward-looking investment decision taking into account both the labor- and marriage-market returns. Second, they enter the marriage market, a frictionless transferable utility setup in the style of Becker (1973), where individual types are defined by their education levels. They decide on which spousal type to get married to or to stay single, and then draw a family composition category (i.e., number of children) according to match-specific empirical probabilities. Third, individuals draw market and home abilities, and households make the dichotomous labor supply choice of whether to work in the market or on home production for each individual. Gender roles are modeled as a disutility that a married couple gets when the wife works in the market (Akerlof

[1] I compare, for example, university-educated married women with one child under the age of 5 and university-educated single women with one child under the age of 5.

[2] As a catch-all term, the norms wedges encapsulate many different factors affecting married women's labor force participation differently from single women's. When I discuss the calibrated values of the norms wedges in section 4.3, I argue how many of these factors fall under the broad concept of homemaker norms for married women. I also address the major disparities unrelated to norms: income effects arising from husband's earnings, tax advantages for married individuals (Guner, Kaygusuz, and Ventura, 2012; Kleven, Kreiner, and Saez, 2009; Blundell, Graber, and Mogstad, 2015; Gayle and Shephard, 2019), and differing child-bearing prospects.

and Kranton, 2000; Bertrand, Kamenica, and Pan, 2015). This disutility factors into the wife's labor supply decision as a "norms wedge" that lowers the value of her market wage. Therefore in my setup, gender roles directly affect labor supply choice, and also indirectly affect marriage and educational choice in anticipation.

I find that married women faced a 44% norms wedge on the market wage in 1940, which declined to 14-15% by 2000 and 2010. If gender norms had stayed at the level of 1940, aggregate market output in 2000 would be lower by 9.0% and aggregate total (market and nonmarket) output would be lower by 5.5%. To gauge the magnitude of these results, I compare them to the output effect of another counterfactual: if married women of 2000 made labor force participation choices based on 1940 wages and home productivities. I find that cultural forces have nearly 80% of the effect of economic forces.

It is the double forces of lower human capital accumulation overall and greater misallocation of the market talent of married women that drive the output effects. First, the average years of schooling of men and women drop by 2.1% and 5.5%, respectively. This is the result of marriage becoming less attractive and thus the marriage market returns to education falling. The marriage rate indeed falls by a third. The labor market returns to education also affects women's education choices, but in the opposite way. Women with the highest education and therefore wages are the first to drop out of marriage, so for these women with a low likelihood of marriage, the likelihood of working actually increases. Women with the lowest education, on the other hand, still have a good chance of getting married, so their likelihood of working falls with more conservative norms. Indeed, married women's labor force participation and cumulative market earnings drop by 34.1% and 28.2%, respectively. In essence, the labor market returns for women get steeper in education. The offsetting effect explains why the drop in schooling is stronger for men than for women. The finding of smaller effects on output than on labor force participation echoes Hsieh et al. (2019)'s findings on the effects of occupation preferences that vary by gender.

The counterfactual also implies that the reduction in the gender norms wedges between 1940 and 2010 accounts for many well-documented trends in the U.S.: a) rising married female labor force participation rate, b) rise in wife's share of household income, c) growing educational attainment of both men and women, and d) increasingly positive selection into marriage by education of both men and women (Case and Deaton, 2017; Juhn and McCue, 2017; Bar et al., 2018).

Since the counterfactual results depend on the model structure, I perform a reduced-form exercise for model validation. For lack of a direct test of model predictions when norms wedges fall, I explore the effects of a shock that *indirectly* affects norms and check that other variables change in the expected direction. Inspired by Fernández et al. (2004), I consider

WW2 draftee casualties as a temporary positive shock to female labor force participation that propagates over the long term through weaker gender norms. Underlying this story is the idea of cultural transmission through exposure (Bisin and Verdier, 2000). I employ a difference-in-difference estimation strategy on the U.S. decennial census comparing high-casualty counties with low-casualty counties in each decade relative to 1940, the last decade before the WW2 shock. The results indicate higher female labor force participation in the high-casualty counties every decade from 1950, but with a spike in 1950, a slight drop in 1960, and a gradual increase over the next decades. At the same time, attitudes become gradually less conservative in the high-casualty counties. Other relevant variables, namely labor force participation by gender and marital status, marriage rate, education, and wages, gradually evolve in a way that is consistent with model predictions when norms wedges fall.

The WW2 reduced-form result also allows me to extend the model to have norms evolving dynamically in response to past female labor force participation. *Economywide* norms evolve in response to *economywide* female labor force participation in the past decade, a relationship that I estimate to match the values of the reduced-form coefficients.[3] The addition of dynamically responding norms enables me to conduct *dynamic* counterfactuals, on how a shock would affect the economy *over time.* A simple policy experiment of closing the gender wage gap in 2010, shows that the economy of 2010 stabilizes to a different equilibrium with higher female labor force participation and lower norms wedges. The exercise illustrates how temporary policies encouraging female labor force participation can have permanent effects.

The analysis in this paper is based on historical data from the United States. Yet, numerous countries in the world are experiencing similar trends as the U.S.: gender attitudes are becoming less conservative,[4] and married women's labor force participation is catching up with single women's. These countries include not only the richer OECD countries but also low- and middle-income countries in Eastern Europe and Latin America. At the same time, one in ten countries of the world still has a lower female labor force participation rate than 1940 U.S. (International Labor Organization, 2019).[5] Thus, this paper informs us about the potential growth consequences and the underlying channels of cultural change in other countries that currently operate under traditional gender roles or are moving away from it.

[3]This extension of the model is compatible with how I identified norms wedges previously, as long as individuals take norms as given and each person does not internalize the effect of their labor supply choice on the norms of future generations as a whole.

[4]Conservativeness in gender attitudes is measured by the fraction agreeing to "When jobs are scarce, men have more right to a job than women," asked in the World Values Survey.

[5]In light of the U-shaped pattern of female labor force participation over the cycle of economic development (Goldin, 1995; Mammen and Paxson, 2000; Ngai, Olivetti, and Petrongolo, 2020), found with the inclusion of unpaid family labor, I should clarify that the mentioned ILO statistic pertains only to paid employment and self-employment.

Contributions to related literature A large literature pioneered by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) study the aggregate implications of various forms of misallocation. A growing number of papers focus on gender differences as a source of misallocation of talent. The most relevant papers are Hsieh et al. (2019), which looks at gender discrimination in the educational and labor markets distorting occupational choice, and Erosa et al. (2022), which studies the gender differences in nonmarket time using married couples only. I add to this literature by focusing primarily on the difference between *married* and *single* women. As I integrate the marriage market matching into the model, I can explore a new set of channels behind the aggregate output implications, such as selection into marriage and marriage market returns to education.

I also contribute to a large body of work that seeks to explain the dramatic rise in married women's labor force participation in the U.S. The explanations proposed thus far can be broadly categorized into two branches: technological progress and cultural change. The first branch includes the invention of birth control pills (Goldin and Katz, 2002), technological advances in housework (Greenwood et al., 2005), and medical progress in pregnancy-related conditions (Albanesi and Olivetti, 2016). The latter, on the other hand, includes changes to divorce laws (Fernández and Wong, 2014), and greater acceptance of working wives by men (Fernández et al., 2004). I add to the second branch by zooming into gender roles associated with marriage, and quantifying within a general equilibrium framework its effect on married women's labor force participation.

My reduced-form analysis around WW2 casualties also speaks to a growing literature on how gender roles change. Kuziemko et al. (2018) explore the birth of the first child as a factor that changes individual's preferences, and Fogli and Veldkmap (2011) and Fernández (2013) model gender roles changing as a result of social learning about the uncertain costs of working. My contribution is to tie the structural model and the reduced-form results together to estimate how norms wedges change in response to past female labor force participation. In addition, I augment the model with this estimated relationship to illustrate how one-off policies can have long-lasting consequences through the dynamic evolution of norms.

Roadmap The rest of the paper flows as follows. Section 2 describes empirical facts that motivate my focus on the distinction between married and single women's labor supply. The structural model in section 3 makes explicit this difference. Section 4 describes the data and the calibration procedure and discusses the calibration results. Section 5 quantifies the effect of changes in gender norms through the lens of the model. Section 6 presents reduced-form results for model validation and a dynamic extension to the model. Section 7 concludes.

2 Motivating Facts

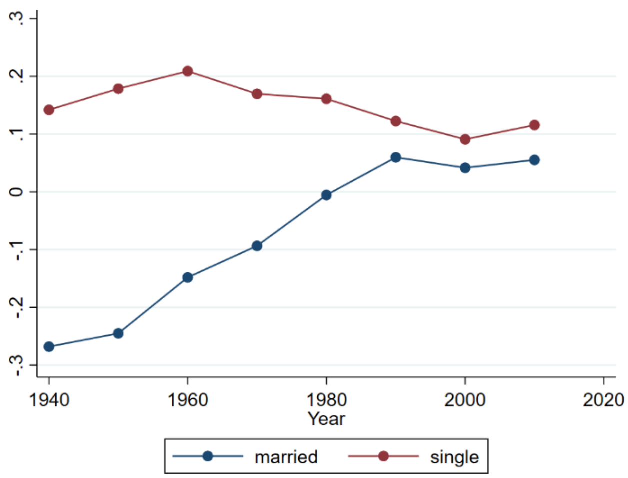
This section presents descriptive facts that motivate my focus on traditional gender roles as the distinguishing feature between married and single women's labor supply, and the evolution of gender roles over time.

"Unexplained" female labor force participation Figure 1 compares the path of the "unexplained" labor force participation (LFP) of married women and single women over time. By "unexplained" LFP, I refer to residuals from the regression of LFP indicator Lab on standard, commonly observed individual characteristics X: age, education, race, and number of children dummies.

$$Lab_{it} = X_{it}\beta + \epsilon_{it}.$$

After running the regression on all females aged 25-54 between 1940 and 2010 in the U.S. decennial census, I take the weighted average of the residuals by marital status and decade.

Figure 1: Residualized Female Labor Force Participation, by Marital Status



Note: This figure compares the labor force participation rates of unmarried (never married, separated, divorced, widowed) to married women between the ages of 25 and 54 in 1940 and 2010, in the United States. The participation rates are residualized for age, education, race, and number of children dummies.

Figure 1 shows that the "unexplained" LFP rose for married women but not for single women. It therefore highlights, firstly, that the LFP choices of married women are very different from single women, and secondly, that this difference shrinks over time. In addition, this difference exists even when LFP status is residualized for the number of children. Therefore, it shows that the distinction between the labor supply behaviors of married and single women extends beyond the presence of children, which has been the dominant factor setting marrieds apart from singles in the literature. The figure also suggests that technological change around child-bearing or child-rearing cannot explain all of this catch-up, leaving room for cultural change around gender roles within marriage as a potential contributor.

Shift in work patterns upon marriage To further corroborate the observation that married women's labor supply choice is disparate from single women's, I study how individuals shift their time use immediately upon marriage.

I follow the event-study approach of Kleven et al. (2019). For this exercise, I use the Panel Survey of Income Dynamics, which is an individual-level panel data where nationally representative individuals of the United States record their weekly paid hours and housework hours.[6] I define event times to be years relative to marriage, such that event time 0 refers to the first year in which an individual's marital status switches to being married from being single. I run the regression

$$housework^g_{ist} = \sum_{j\neq -1} \alpha^g_j \cdot \mathbb{1}(j=t) + \sum_k \beta^g_k \cdot \mathbb{1}(k = age_{is}) + \sum_y \gamma^g_y \cdot \mathbb{1}(y=s) + \nu^g_{ist}$$

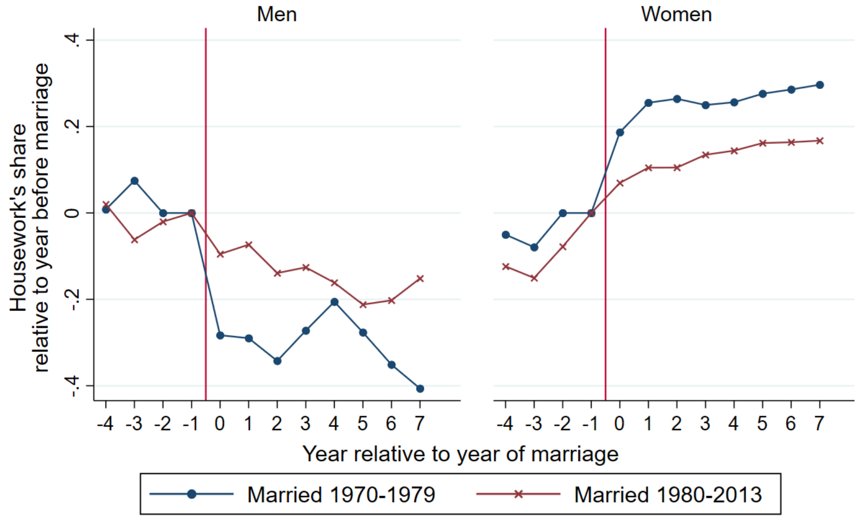
where $housework^g_{ist}$ denotes the housework's share of total work (market work and housework) hours of individual i of gender g in year s at event time t. This regression tracks how *housework* changes as a function of event time, while controlling for life-cycle dynamics via the age dummies and time trends via the year dummies. The event time coefficients ($\hat{\alpha}^g_t$) are then normalized by $\mathbb{E}[\tilde{Y}^g_{ist}|t]$, where $\tilde{Y}^g_{ist} \equiv \sum_k \hat{\beta}^g_k \cdot \mathbb{1}(k = age_{is}) + \sum_y \hat{\gamma}^g_y \cdot \mathbb{1}(y=s)$ is the predicted outcome when excluding the effect of the event time. $\hat{\alpha}^g_t / \mathbb{E}[\tilde{Y}^g_{ist}|t]$ are plotted in Figure 2.[7]

The blue lines of Figure 2 illustrate a sharp jump in women's share of housework hours among total work hours and a sharp drop in men's, immediately upon transition from singlehood to marriage, for marriages in the 1970s. This finding highlights that being married shifts the responsibility of house chores to women. The magnitude of the time use shift is sizable. The jump in the housework's share for women amounts to about half of the jump associated with the birth of the first child for the same sample of women. As the sample consists of couples who had no childbirths in the first three years of marriage, the figure additionally suggests that marriage itself – independent of the presence of children – subjects women to gender roles. This idea resonates with married women's LFP falling short of single women's even when residualized for the number of children, shown in Figure 1.

[6] The exact survey question is "About how much time does (he/she) (do you) spend on this housework in an average week–I mean time spent cooking, cleaning, and other work around the house?"

[7] The actual event time coefficients (α^g_j) are plotted in Figure A2 of the Appendix. They are statistically insignificant at the 5% level prior to marriage, and are significantly negative for men and significantly positive for women post-marriage

Figure 2: Housework's Share of Housework and Market Hours, Among Couples whose First Child is Born $\geq$ 4 years After Marriage



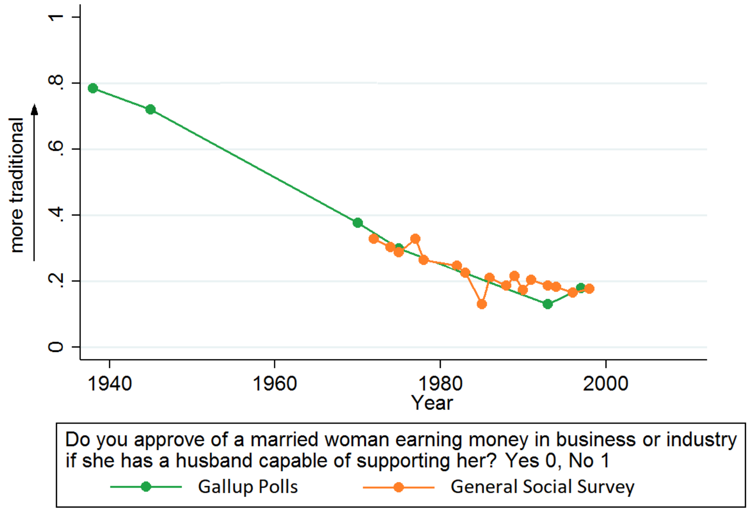
Note: This figure plots $\hat{\alpha}^g_t/\mathbb{E}[\tilde{Y}^g_{ist}|t]$, the change in the share of housework among total work (market work and housework) hours by gender, around the year of marriage. The red vertical line plots the timing of marriage. Individuals are unmarried household heads without any live-in partners in the years to the left of the red line, and they are married with live-in spouses in the years to the right of the red line.

The red lines of Figure 2, on the other hand, demonstrate that there are no sharp changes in time allocation for marriages that take place later. The event study coefficients $\hat{\alpha}^g_t$ are also statistically insignificant around the year of marriage. I take this null effect for later marriages as suggestive of the decline over time in the division of labor according to traditional roles.

Attitudinal survey trends The last motivating fact illustrates the weakening of gender roles of married women over time, which has been previously documented (e.g. Fernández, 2013; Goussé, Jacquemet, and Robin, 2017). Among various survey questions on gender attitudes in the ROPER polls database that I identified from a keyword search of "women" and "gender," the question that was asked for by far the longest period was whether one approved of a married woman working if she had a husband capable of supporting her. Figure 3 shows that while close to 80% answered "No" in 1938, in 1998 less than 20% did so. The survey question ceased to be asked afterward, which itself could be suggestive of the question being less controversial and thus of less interest than before.

Other survey questions on gender roles of married individuals have been asked after 1998, however, with answers confirming a continual (but slowing) trend towards less traditional attitudes. These trends are shown in Figure A1 of the Appendix. Answers to these questions are used in section 4.3 to cross-check the validity of the model-derived measures of norms.

Figure 3: Trend in Attitudinal Survey Answers



Note: This figure plots the fraction of respondents disapproving of a married woman working if she has a husband capable of supporting her, according to the Gallup Polls and the General Social Survey (GSS). The two surveys asked nearly identical questions. The Gallup Polls' specific question was "Should a married woman earn money if she has a husband capable of supporting her?", while the GSS asked, "Do you approve of a married women earning money in business or industry if she has a husband capable of supporting her?"

3 Model

The motivating facts of Section 2 suggest that married women's labor supply decisions are different from single women's, that gender roles could be one of the factors driving this gap, and that this gap is shrinking over time. The model developed in this section formalizes the difference in married and single women's labor supply decisions.

Timing The decisions of individuals are divided into three stages. In the first stage, individuals choose their level of education as a forward-looking investment, balancing the returns to education in the labor and marriage markets with the cost of education (Chiappori, Iyigun, and Weiss, 2009). In the second stage, they enter a frictionless transferable utility (TU) marriage market (Shapley and Shubik, 1971; Becker, 1973), where "types" of individuals equal their education levels chosen previously, and decide on which spousal type to get married to or to stay single. The resulting match is thus characterized by the education levels of the husband and wife for married couples and by one's own education level for singles. Households then exogenously get assigned family composition categories based on the number of children under the age of 5 and under the age of 18 in the household, according to match-specific empirical probabilities. Individuals subsequently enter the third and last stage, each characterized by group: the tuple of (gender, marriage match, family composition). After individuals draw idiosyncratic market and home abilities, households make the dichotomous labor supply choice of whether to work in the market or in home production for each individual, taking as given a) the group-specific market wage that the representative firm of the economy pays, b) the group-specific value of home production, and c) the group-specific disutility that a married couple gets upon deviation from traditional gender roles, i.e. when the wife works in the market. After the labor supply decisions are made, households consume and realize utilities.

Since I solve the model backwards, I describe each stage in greater detail, in reverse order.

3.1 Economic utilities and optimal labor supply choices

Individual utilities consist of an economic and a predetermined noneconomic component. The *economic* component is characterized by the utility functions below, adapted from Chiappori, Costa-Dias, and Meghir (2018).[8] The economic gains from marriage arise from two sources. Firstly, there are economies of scale generated by the consumption of public goods. Secondly, marriage enables risk sharing between the two spouses against uncertain future public and private consumption.

Married Individuals

Consider a married household composed of husband m and wife f. Individual $i \in \{m, f\}$ gets the following utility:

$$u_i(Q, C_i, L_f) = \ln(Q) + \ln(C_i - \tau_i w_f L_f) \tag{1}$$

where C_i is the consumption of the private good, and Q is the consumption of the public good (e.g. expenditures on housing, children, heating). $L_i \in \{0, 1\}$ denotes labor market participation and w_i market wage, of individual i. Critically, married individual i gets disutility when the wife works in the market. Parameter τ_i measures the value of the disutility as a proportion of the market income brought home by the wife.

From an ordinal perspective, this utility belongs to Bergstrom and Cornes (1983)'s Generalized Quasi Linear (GQL) family. Hence, for any realization of family income, it satisfies the transferable utility (TU) property. Under TU, utility can be transferred between spouses at a fixed rate of exchange, and so for Pareto-efficiency, a couple acts as a single decision unit that maximizes the joint marital output. As the set of Pareto efficient allocations is an ordinal concept, any cardinalization of u can be used for the definition of joint marital out-

[8] A more general formulation of the utility function that is consistent with my model is described in Appendix C.1.

put. I use $\exp u_i$ as the cardinalization of i's preferences. Then, conditional on the couple's labor choices, any Pareto efficient allocation maximizes the sum of the spouses' exponential utilities, $\exp u_m + \exp u_f$.

Therefore, conditional on labor market participation choices, a married couple solves

$$\begin{aligned} &\max_{Q,C} Q(C - \tau w_f L_f) \\ \text{s.t.} \quad & pQ + C = w_m L_m + w_f L_f + h_m(1 - L_m) + h_f(1 - L_f) \end{aligned} \tag{2}$$

$C \equiv C_m + C_f$ denotes total expenditure on private goods, $\tau \equiv \tau_m + \tau_f$ is the couple's joint disutility from the wife working in the market, and p is the price of the public good relative to the private good (the numeraire). Moreover, h_i refers to the home productivity of individual i.[9] The solutions to the maximization problem given by (2) are

$$\begin{aligned} Q &= \frac{w_m L_m + (1-\tau) w_f L_f + h_m(1 - L_m) + h_f(1 - L_f)}{2p} \\ C &= pQ + \tau w_f L_f \end{aligned} \tag{3}$$

Let us describe the intra-household allocation, before market earnings and home production values are realized. Efficient sharing of the risks against the uncertainty of earnings and home productivities implies that the ratio of marginal utilities of private consumption is constant and equal to the Pareto weight (μ),[10] which is endogenously determined in the marriage market:

$$\frac{\partial u_m}{\partial C_m} = \mu \frac{\partial u_f}{\partial C_f}$$

The resulting indirect utilities are:

$$v_m = 2 \ln Q + \ln p + \ln \frac{1}{1+\mu}, \qquad v_f = 2 \ln Q + \ln p + \ln \frac{\mu}{1+\mu} \tag{4}$$

The total economic utility generated from this marriage then is

$$v = v_m + v_f = \breve{v} + \ln \frac{\mu}{(1+\mu)^2} \tag{5}$$

where $\breve{v} \equiv 4 \ln Q + 2 \ln p$. It is straightforward that the farther the wife's Pareto weight μ is from 1, the husband's, the smaller the total economic utility of the couple.

[9] It might seem odd that according to the budget constraint, private and public goods can be bought with "income" from home production. However, Appendix section C.2 shows that the utility maximization problem remains identical if I divide goods into market goods and home-produced goods, and have two separate budget constraints for each.

[10] Browning, Chiappori, and Weiss (2014) characterize intra-household allocations under efficient risk-sharing.

Clearly, the couple makes labor choices to maximize Q.[11] Hence, from (3), the optimal labor choices are

$$\begin{aligned} L_m^* &= \mathbb{1}\left[w_m \geq h_m\right] \\ L_f^* &= \mathbb{1}\left[(1-\tau)w_f \geq h_f\right] \end{aligned} \tag{6}$$

In the optimal labor choice of married women, τ enters as a "norms wedge." In deciding her labor supply, a married woman values her market wage lower than its face value, as if it is taxed.[12] One feature that stands out in solution (6) is the independence of the husband's and wife's labor supply choices. While this feature makes studying the selection into working in the labor market by either the husband or the wife easy, it is not realistic. I consider an extension of the model in Appendix section C.4: assuming men always work, I relax the independence so that the labor supply decisions of women are responsive to their husbands' market earnings. I show later in section 4.3 that the benchmark model with independent spousal decisions performs well up to 2000.

The flexibility of the model allows for another extension (Appendix section C.5) where norms wedges apply also to married *men*, where the nontraditional activity for them is home production.[13]

Single Individuals

To distinguish from the married case, I use the hat symbol for singles. The economic utilities of singles follow the same formulation as for married couples, except they are not subject to gender roles and hence single women do not receive disutilities from working in the labor market.

A single individual i solves

$$\begin{aligned} &\max_{Q_i, C_i, L_i} \hat{u}_i(\hat{Q}_i, \hat{C}_i) = \ln(\hat{Q}_i) + \ln(\hat{C}_i) \\ &\text{s.t.} \qquad p\hat{Q}_i + \hat{C}_i = w_i \hat{L}_i + h_i(1-\hat{L}_i) \end{aligned} \tag{7}$$

The resulting indirect utility is

$$\hat{v}_i = 2\ln\hat{Q}_i + \ln p \tag{8}$$

[11]Moreover, as $\frac{\partial^2 \ln Q}{\partial w_m \partial w_f} = 0$, the model does not predict any assortative matching on market earnings.

[12]The norms wedge functions as a catch-all term that encapsulates many different factors that make married women's LFP diverge from single women's LFP besides the wage differentials. See section 4.3 for a discussion of what these factors include, and how they fall under the broad concept of homemaker norms for married women.

[13]Appendix section E.1 presents the calibration of the "breadwinner" norms wedges for married men. These decrease over time, much like the female counterpart. However, they come with some caveats, which are discussed.

and the optimal labor choice that maximizes $\hat{Q}_i$ is

$$\hat{L}_i^* = \mathbb{1}\left[w_i \geq h_i\right] \tag{9}$$

Market Income and Home Production Value As optimal labor supply choices for married and single individuals depend on the comparison of market earnings to home production value, it is imperative to discuss how they are determined. Market income and home production value depend on idiosyncratic market and home abilities, as well as components common to groups defined by gender, marriage match, and family composition. Note that after the marriage matching stage, family composition is given exogenously according to match-specific empirical probabilities. Hence, the probability that a (husband education level q, wife education level r) match has a family composition $\mathcal{K}$, denoted as $d^{qr}(\mathcal{K})$, is simply found from the data.[14]

An individual i of gender g in a (q, r) match, with family composition $\mathcal{K}$, receives income

$$w_i = \bar{w}_g^{qr}(\mathcal{K})\epsilon_i^w, \quad g \in \{M, F\}$$

where $\bar{w}_g^{qr}(\mathcal{K})$ is the market income per unit of effective labor for i's group, and ϵ_i^w is i's market ability. The reason why group wages differ can be thought of as a combination of selection and treatment effects. For instance, I am flexibly letting married women have different market productivity from single women because individuals who get married might be different from those who are single (selection), and marriage might causally affect market productivity (treatment effect). Similarly, college-educated women married to high-school dropout husbands are allowed to have different wages from college-educated women married to college-educated husbands as a result of both selection and treatment effects.

Where i's education level equals s, i's home production value is given by

$$h_i = \bar{h}_g^s(\mathcal{K})\epsilon_i^h, \quad g \in \{M, F\}$$

The group component of home production value, $\bar{h}_g^s(\mathcal{K})$, varies by gender, own education level, and family composition. Inherent in this assumption is that marital status and spousal type does not matter for home productivity, which is necessary for me to be able to later disentangle norm wedge parameters from home productivity parameters.

I assume that market abilities ϵ^w and home production abilites ϵ^h are drawn independently and identically from the Fréchet distribution with shape parameter θ, after the marriage

[14] It is possible for singles to have children in my model, because singles include never-married, divorced, separated, and widowed individuals. This grouping of singles is equivalent to assuming that divorces, separations, and widowhoods occur via shocks exogenous to the schooling years of the couple.

matching stage.[15] The cumulative distribution functions for these abilities are

$$F(\epsilon^w) = F(\epsilon^h) = F(x) = \exp\left\{-x^{-\theta}\right\}.$$

From the convenient property of Fréchet distributions, the probability that a woman in a (q, r) match with family composition $\mathcal{K}$ works in the labor market is:

$$P_F^{qr}(\mathcal{K}) \equiv \mathbb{P}\left((1-\tau^{qr}(\mathcal{K}))\bar{w}_F^{qr}(\mathcal{K})\epsilon^w > \bar{h}_F^r(\mathcal{K})\epsilon^h\right) = \frac{[(1-\tau^{qr}(\mathcal{K}))\bar{w}_F^{qr}(\mathcal{K})]^\theta}{[(1-\tau^{qr}(\mathcal{K}))\bar{w}_F^{qr}(\mathcal{K})]^\theta + [\bar{h}_F^r(\mathcal{K})]^\theta} \quad (10)$$

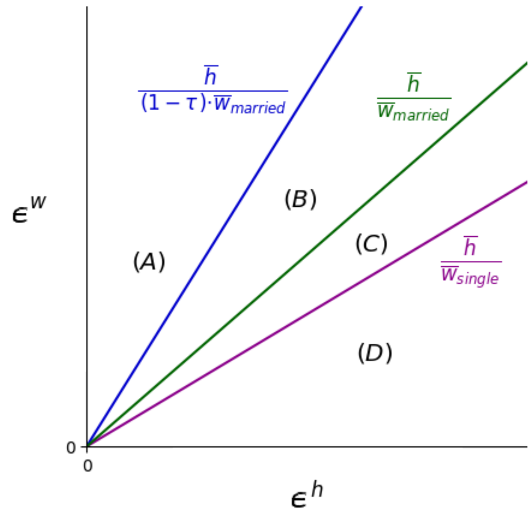
The maximum likelihood estimator for this probability is the labor force participation rate of the women in this group. Equation (10) is useful for calibrating parameters later in section 4. Moreover, Figure 4 illustrates how sorting across market work and home production by market and home abilities occurs for married and single women. Another implication of Fréchet abilities useful for calibration later on is that the average wage of the women working in the market is[16] :

$$\begin{aligned} \text{avgwage}_F^{qr}(\mathcal{K}) &= \bar{w}_F^{qr}(\mathcal{K})\mathbb{E}\left[\epsilon^w | (1-\tau^{qr}(\mathcal{K}))\bar{w}_F^{qr}(\mathcal{K})\epsilon^w > \bar{h}_F^r(\mathcal{K})\epsilon^h\right] \\ &= \bar{w}_F^{qr}(\mathcal{K})\left(\frac{1}{P_F^{qr}(\mathcal{K})}\right)^{\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}\right) \end{aligned} \quad (11)$$

[15]The extensive literature on returns to schooling (Griliches, 1977) highlights the correlation between schooling and unobserved abilities. Hence, it might be more plausible that the market and home production abilities are drawn from education-specific distributions. However, in incorporating the correlation between schooling and unobserved abilities into the model, I take a shortcut by assuming that different education levels result in abilities being drawn from the same distribution scaled by different constants. In other words, where ϵ_s^w is the market ability drawn from a distribution specific to education level s, $\epsilon_s^w = \alpha^s \epsilon^w$. Then, it is possible to take the scaling constants (α^s) out of intrinsic abilities and have schooling-specific wages and home productivities incorporate the scaling constants.

[16]Where ϵ^{w*} is the market ability ϵ^w conditional on working in the market, its cumulative distribution function is $F^*(x) = \exp\left\{-\frac{1}{P}x^{-\theta}\right\}$ where P is the fraction working in the market. In other words, F^* follows the Fréchet distribution with shape parameter θ and scale parameter $\left(\frac{1}{P}\right)^{\frac{1}{\theta}}$.

Figure 4: Sorting Across Market Work and Home Production of Married and Single Women



Note: This figure plots how labor allocation between market work and home production is determined for different combinations of market abilities ϵ^w and home abilities ϵ^h, for married and single women with the same education level and family composition (in simplified notation).
(A): For a married woman to work in the market, she must be very talented in market work.
(B): If the norms wedge was removed, more married women would engage in market work.
(C): Along with (A) and (B), the ability combinations of single women doing market work.
(D): Single women who work at home are very talented in home production.

3.2 Marriage market

Marriage matching occurs based on the expected value of the economic utilities delineated in the previous subsection, together with the noneconomic utilities, of each match.

Following Choo and Siow (2006)'s matching under transferable utility (TU) with random preferences, consider an economy consisting of S types of men and women. These types are defined by the level of education determined prior to the matching stage. Denote n^{qr} as the number of marriages between type-q men and type-r women, n^{q0} as the number of single type-q men, and n^{0r} as the number of single type-r women. Also, M^q is the number of type-q men, and F^r the number of type-r women. The following accounting identities must hold:

$$n^{q0} + \sum_{r=1}^{S} n^{qr} = M^q \qquad \forall\, q = 1, ..., S \tag{12}$$

$$n^{0r} + \sum_{q=1}^{S} n^{qr} = F^r \qquad \forall\, r = 1, ..., S \tag{13}$$

In this TU model, a type-q man must transfer π^{qr} amount of utility to a type-r woman to marry her. The utility of type-q man m marrying a type-r woman at time t is

$$V_m^{qr} = \mathbb{E}(v_m^{qr}) - \pi^{qr} + \psi^{qr} + \varepsilon_m^{qr}$$

where $\mathbb{E}(v_m^{qr})$ is the expected economic utility of man m married to a type-r woman, ψ^{qr} is the noneconomic utility ("marital bliss") enjoyed by the couple, and ε_m^{qr} is m's random preference for the match drawn independently and identically from the type I extreme-value distribution. Let $r = 0$ denote the case of singlehood, with $v_m^{q0} \equiv \hat{v}_m^q$, $\tau^{q0} = 0$, and $\psi^{q0} = 0$.

Similarly, the utility of type-r woman f marrying a type-q man is

$$V_f^{qr} = \mathbb{E}(v_f^{qr}) + \pi^{qr} + \psi^{qr} + \varepsilon_f^{qr}$$

where ε_f^{qr} is f's random preference for the match drawn independently and identically from the type I extreme-value distribution.

Combining the marriage market-clearing condition[17] and equation (5) gives

$$\frac{n^{qr}}{\sqrt{n^{q0}n^{0r}}} = \frac{\mathbb{E}(\breve{v}^{qr}) - \mathbb{E}(\hat{v}_m^q) - \mathbb{E}(\hat{v}_f^r)}{2} + \Psi^{qr} \tag{14}$$

where $\Psi^{qr} \equiv \psi^{qr} + \frac{\ln \mu^{qr} - 2\ln(1+\mu^{qr})}{2}$. The first term on the right-hand side of equation (14) is the gain from marriage relative to singlehood, from the couple being able to enjoy a greater consumption of the public good together. Ψ^{qr} signifies the utility from marital bliss and the utility from the intra-household allocation of resources based on Pareto weights.[18]

3.3 Education

In this section, I describe the women's educational choice problem, without loss of generality. Woman f chooses the education level with the maximum expected utility:

$$\max_{r=1,\ldots,S} U_F^r$$

where

$$U_F^r = \sum_{q=0}^{S} \left[\frac{n^{qr}}{F^r} \Big(\mathbb{E}(v_f^{qr}) + \pi^{qr} + \psi^{qr} \Big) \right] - c_F^r - \xi^r \tag{15}$$

[17]See Appendix section C.6 for greater details on the derivation of the marriage market equilibrium.

[18]ψ^{qr} and μ^{qr} cannot be separately identified. As can be seen in Appendix section D.3, μ^{qr} would only be identified if the equilibrium transfers in the marriage market were observable, but they are not. Hence, I seek to identify Ψ^{qr}. Identifying Ψ^{qr} is sufficient for running counterfactuals.

Individuals are forward-looking. The expected utility from schooling level r depends on the consequent matching probabilities in the marriage market and the expected utility in each type of match. The costs of schooling level r, on the other hand, consist of the gender-specific direct utility cost c_F^r and idiosyncratic cost ξ^r, drawn independently and identically from the Type I extreme value distribution.

The distribution of the idiosyncratic schooling costs imply that the probability an individual of gender g chooses schooling level s is

$$\mathbb{P}(s = \arg\max_{s'=1,..,S} U_g^{s'}) = \frac{\exp\{U_g^s\}}{\sum_{s'=1}^S \exp\{U_g^{s'}\}}$$

The maximum likelihood estimator of this probability is $\frac{F^s}{\sum_{s'=1}^S F^{s'}}$ for women and $\frac{M^s}{\sum_{s'=1}^S M^{s'}}$ for men, i.e. the shares of individuals with education level s for each gender.

3.4 Firms

A representative firm produces market output Y^{mkt}. Although there are two market goods in this model, the private good and the public good, I assume that they are derived from the same market output. The relative price p merely measures how much more market output is needed for 1 unit of public good, relative to 1 unit of private good. This simplification is innocuous, given that the value of p has no consequence for equilibrium education, marriage, and labor decisions.

I assume the most simplistic set-up on the firm's side. The firm's production function is linear in male and female effective labor, $\mathcal{M}$ and $\mathcal{F}$:

$$Y^{mkt} = B(\mathcal{M} + \mathcal{F}) \tag{16}$$

Normalize, as 1 unit of effective labor, the labor provided by single males with schooling level of 1, market ability of 1, and zero children ($\mathcal{K}_0$), i.e. $\bar{w}_M^{10}(\mathcal{K}_0)$.

$$\mathcal{M} = \sum_{q=1}^{S}\sum_{r=0}^{S}\sum_{\mathcal{K}} n^{qr} d^{qr}(\mathcal{K}) \left(\frac{\bar{w}_M^{qr}(\mathcal{K})}{\bar{w}_M^{10}(\mathcal{K}_0)}\right) \left(P_M^{qr}(\mathcal{K})\right)^{1-\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right)$$

$$\mathcal{F} = \sum_{r=1}^{S}\sum_{q=0}^{S}\sum_{\mathcal{K}} n^{qr} d^{qr}(\mathcal{K}) \left(\frac{\bar{w}_F^{qr}(\mathcal{K})}{\bar{w}_M^{10}(\mathcal{K}_0)}\right) \left(P_F^{qr}(\mathcal{K})\right)^{1-\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right)$$

3.5 Aggregate output

Aggregate output is a combination of market output and home production:

$$Y = Y^{mkt} + Y^{home}$$

where

$$Y^{home} = B \sum_{g \in \{M,F\}} \sum_{(q,r)} \sum_{\mathcal{K}} n^{qr} d^{qr}(\mathcal{K}) \left(\frac{\bar{h}_g^r(\mathcal{K})}{\bar{w}_M^{10}(\mathcal{K}_0)} \right) \left(1 - P_g^{qr}(\mathcal{K})\right)^{1-\frac{1}{\theta}} \Gamma \left(1 - \frac{1}{\theta}\right) \tag{17}$$

3.6 Equilibrium

An equilibrium in this economy consists of schooling choice q for a man, schooling choice r for a woman, marital transfers π^{qr}, marriage matches (q, r), public consumption Q, private consumption C_i, labor market participation L_i, total efficient male labor $\mathcal{M}$, total efficient female labor $\mathcal{F}$, market wage, market output Y^{mkt}, total home production Y^{home}, and aggregate output Y, such that

1. Individuals choose the schooling level offering the greatest expected utility, taking as given the probability of resulting in a particular match and the expected utility from that match.
2. After schooling choices are made, equilibrium marital transfers $\{\pi^{qr}\}$ equate the supply and demand for each marriage match (q, r) based on the expected utility from each match.
3. After the matching stage and exogenous determination of family composition, each individual chooses public good consumption Q, private good consumption C_i, and labor supply L_i to maximize their utility function. The individual maximizes equation (1) jointly with their spouse if married and maximizes equation (7) independently if single.
4. A representative firm hires effective male labor $\mathcal{M}$ and effective female labor $\mathcal{F}$, and pays wage equal to the technology parameter B in equation (16).
5. Market output Y^{mkt} is given by equation (16), and total home production Y^{home} by (17).
6. Aggregate output of the economy Y is given by the sum of Y^{mkt} and Y^{home}.

3.7 Intuition for aggregate output effects of norms wedges

In the model, how is aggregate output affected by changes in norms wedges? When the norms wedge on market wage for married women decreases, there can be aggregate output effects arising from each of the three stages (in reverse order) of labor supply, marriage, and education choices for women.

First, at the labor supply stage, sorting across market work and home production of *married* women is more aligned with productivity. This channel increases aggregate output.

Second, at the marriage matching stage, marriage becomes more attractive as the disutility from wives working in the market is lower. Then some women who would otherwise have been single would now be married and therefore be newly subject to the norms wedge. As the norms wedge prevents some of these women from pursuing their comparative advantage, aggregate output is lower. There is another effect occurring at the matching stage. The women who are newly induced to be married now receive married wages. As whether married wages are higher or lower than single wages is an empirical question, this channel has an ambiguous effect on aggregate output.

Third, at the educational choice stage, if there is positive assortative matching on education then the greater likelihood of marriage increases the young women's marriage market incentive to get educated. In addition, if education is more effective in increasing market productivity than home productivity, the labor market incentive to get educated depends on how the expected likelihood of working changes. The likelihood of working in the market depends on the balance between i) the increased likelihood of being newly subject to the norms wedges as marriage rates go up, and ii) the norms wedges when married going down. How the labor market returns to education changes is an empirical question, but if education does increase on net, then aggregate output would also increase through higher market wages and home productivities.

Overall, the effect on aggregate output would depend on the parameter values.

4 Data and Parameter Inference

4.1 Data

To simulate the U.S. economy within the model framework, I use the U.S. decennial census, consisting of 1-in-100 national random sample of individuals. The nice feature of the U.S. census is that data is collected on all household members so that labor market information is available for both spouses among married couples. Because the presence of other income-earning household members may perturb individual labor decisions, I restrict the sample to either household heads or spouses of heads. I further restrict the sample to individuals aged between 25 and 54, after education is complete and when individuals are the most economically active.[19]

The model in section 3 is fitted to the census data every decade, assuming that the data is a reflection of the model equilibrium. By calibrating the model separately by decade, I am allowing all model parameters to change flexibly over time, including family composition probabilities $\{d^{qr}(\mathcal{K})\}$, group market wages $\{\bar{w}_M^{qr}(\mathcal{K}), \bar{w}_F^{qr}(\mathcal{K})\}$ and home productivities $\{\bar{h}_M^{qr}(\mathcal{K}), \bar{h}_F^{r}(\mathcal{K})\}$, the inverse measure of the dispersion of market and home abilities $\{\theta\}$, gender norms wedges $\{\tau^{qr}(\mathcal{K})\}$, the expected utility enjoyed by a (q,r) match $\{\Psi^{qr}\}$, and the cost of each schooling level $\{C_M^q, C_F^r\}$.

The practical advantage of my model set-up is that the model is not demanding on the data; the variables needed for these parameters to be inferred are market wage, labor force participation status, marital status, education, and children. As the earliest decade in which all these variables are observed is 1940, I use the decennial census from 1940 to 2010. The census in 1950 is not used, however, because the 1950 data does not include information on spousal wages.

4.2 Steps for Parameter Inference

1. $d^{qr}(\mathcal{K})$: probability of a (q,r) match having family composition $\mathcal{K}$

   Set at the empirical probabilities.

2. θ: inverse measure of dispersion of market and home abilities

   Making use of the fact that wages of individuals working in the market follow a Fréchet distribution, I estimate θ through maximum likelihood. Where x_n is the market ability of observation n and P_n denotes the fraction of workers in observation n's group, the maximum likelihood estimator for θ is:[20]

$$\tilde{\theta}_{MLE} = \arg\max_{\theta\in(0,\infty)} \sum_{n=1}^{Nobs} \left[\ln\theta - \ln P_n - x_n^{-\theta} P_n^{-1} - (\theta+1)\ln x_n\right]$$

3. $\bar{w}_g^{qr}(\mathcal{K})$: group market productivity per unit of effective labor

   Using the estimate of θ found in step 2 and the average wage and proportion of market-workers in each group in the data, I can back out $\bar{w}_g^{qr}(\mathcal{K})$ from equation (11).

$$\bar{w}_g^{qr}(\mathcal{K}) = \text{avgwage}_g^{qr}(\mathcal{K}) \left(P_g^{qr}(\mathcal{K})\right)^{\frac{1}{\theta}} \frac{1}{\Gamma(1-1/\theta)}$$

[19] Appendix figure A3 shows that the age range of 25-54 is the most economically active 30-year window.

[20] How I derive the likelihood function and how I isolate market abilities from observed market wages are detailed in Appendix sections D.1 and D.2, respectively.

4. $\bar{h}^s_g(\mathcal{K})$: group home productivity per unit of effective labor

   $\bar{h}^r_F(\mathcal{K})$ (similarly, $\bar{h}^q_M(\mathcal{K})$) can be backed out from equation (10), armed with θ found in step 2 and the average wage and proportion of market-workers among *single* women with r years of schooling and family composition $\mathcal{K}$.

$$\bar{h}^r_F(\mathcal{K}) = \text{avgwage}^{0r}_F(\mathcal{K})\left(1 - P^{0r}_F(\mathcal{K})\right)^{\frac{1}{\theta}} \frac{1}{\Gamma(1-1/\theta)}$$

5. $\tau^{qr}(\mathcal{K})$: group norms wedges

   $\tau^{qr}(\mathcal{K})$ is backed out from equation (10) using $\bar{w}^{qr}_F(\mathcal{K})$, $\bar{h}^r_F(\mathcal{K})$, and the fraction of market workers in a group of married women. The idea is that norms wedges are high if the fraction working in the market is much lower than is predicted from market and home productivities.

$$\tau^{qr}(\mathcal{K}) = 1 - \frac{\text{avgwage}^{0r}_F(\mathcal{K})}{\text{avgwage}^{qr}_F(\mathcal{K})}\left(\frac{1 - P^{0r}_F(\mathcal{K})}{1 - P^{qr}_F(\mathcal{K})}\right)^{\frac{1}{\theta}} \tag{18}$$

   Intuitively, the disutility from wives working is inferred by comparing the labor choices of married and single women sharing the same level of education and the same family composition. The difference in their labor market participation rates that cannot be explained by wage differentials is attributed to gender norms.

6. Ψ^{qr}: utility from marital bliss and intra-household resource allocation, in a (q, r) match

   From equation (14),

$$\Psi^{qr} = \frac{n^{qr}}{\sqrt{n^{q0}n^{0r}}} - 2A^{qr} + \hat{A}^q_M + \hat{A}^r_F. \tag{19}$$

   where

$$A^{qr} = \sum_{\mathcal{K}} d^{qr}(\mathcal{K})\mathbb{E}\Big[\ln\Big(\bar{w}^{qr}_M(\mathcal{K})\epsilon^w_m L^*_m + \bar{h}^{qr}_M(\mathcal{K})\epsilon^h_m(1 - L^*_m) + \\ \big[1 - \tau^{qr}(\mathcal{K})\big]\bar{w}^{qr}_F(\mathcal{K})\epsilon^w_f L^*_f + \bar{h}^r_F(\mathcal{K})\epsilon^h_f(1 - L^*_f)\Big)\Big]$$

   and

$$\hat{A}^s_g = \sum_{\mathcal{K}} d^{qr}(\mathcal{K})\mathbb{E}\Big[\ln\Big(\bar{w}^{qr}_g(\mathcal{K})\epsilon^w_i \hat{L}^*_i + \bar{h}^s_g(\mathcal{K})\epsilon^h_i(1 - \hat{L}^*_i)\Big)\Big] \quad (q = 0, r = s \text{ if } g = F, \text{ and } q = s, r = 0 \text{ if } g = M)$$

   There are no closed-form expressions for A^{qr} or $\hat{A}^s_g$ so I simulate them to back out Ψ^{qr}.[21]

[21] To permit a more realistic curvature of the utility function, I use the CRRA utility function instead of log utility in simulating A^{qr} and $\hat{A}^s_g$. Note that the CRRA utility fully conforms with the set of conditions of the general utility function, described in Appendix section C.1. The coefficient of relative risk aversion

7. c_g^s: utility cost of aquiring education s

Use equation (15). $\{c_F^r\}_{r=1,...,S}$ are found as the solution to the system of equations

$$\frac{F^r}{\sum_{r'=1}^{S} F^{r'}} = \frac{\exp\left\{\sum_{q=0}^{S} \frac{n^{qr}}{F^r} H^{qr} - c_F^r\right\}}{\sum_{r'=1}^{S} \exp\left\{\sum_{q=0}^{S} \frac{n^{qr'}}{F^{r'}} H^{qr'} - c_F^{r'}\right\}} \qquad \forall\, r = 1, ..., S$$

where

$$H^{qr} = \begin{cases} 2\hat{A}_F^r & \text{if } q = 0 \\ 2A^{qr} + \Psi^{qr} + \frac{\ln n^{q0} - \ln n^{0r}}{2} - \hat{A}_M^q + \hat{A}_F^r & \text{if } q = 1, ..., S \end{cases}$$

Additionally, $\{c_M^q\}_{q=1,...,S}$ are found similarly.

4.3 Calibration Results and Discussion

I now provide a discussion of the calibrated parameter values, and how they match similar estimates in the literature, related measures from external data sources, or well-documented stylized facts.

$\boldsymbol{\theta}$: inverse measure of dispersion of market and home abilities

As shown in Table 1, the estimates of θ range from 1.55 to 2.08, which is similar to Hsieh et al. (2019)'s estimate of 1.52 for the Fréchet shape parameter dictating the dispersion of

Table 1: Maximum Likelihood Estimate of θ

	1940	1960	1970	1980	1990	2000	2010
$\hat{\theta}$	1.55***	1.84***	1.85***	1.72***	2.05***	2.08***	2.01***
	(18.0)	(28.3)	(19.3)	(13.0)	(21.9)	(23.1)	(45.9)
N	156404	314751	315418	560886	697820	791308	726419

Note: t statistics using standard errors clustered by sex in parentheses, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$
When the standard errors are not clustered, following the model assumption of independently drawn market abilities, the t statistic is incredibly large due to the large sample size.
See step 2 of section 4.2 for the maximum likelihood estimation strategy.

used is 0.71, which Chetty (2006) finds to be the average value across 33 different estimation methods spanning structural lifecycle models, natural experiments, earned income responses, and macroeconomic trends. Sensitivity checks to the value of the coefficient of relative risk aversion show that the result of the main counterfactual — the effect on the total output of the 2000 economy of norms wedges reverting to 1940 levels — is -7.48% if the coefficient is 0.61, -5.48% if 0.71 (preferred value), and -5.46% if 0.81.

abilities across occupations. It is also close to their choice to use 2 for conducting counterfactuals. Moreover, Appendix figure A4 visually shows that the distribution of inferred market abilities closely resembles the probability distribution of the Fréchet distribution, supporting the assumption of Fréchet-distributed abilities.

$\bar{w}, \bar{h}$: group market and home productivity

To calibrate the group-specific market and home productivities, I first need to specify the groups. Each group is defined by gender, schooling pair, and family composition. I must ensure that each group is large enough since I match population moments to sample analogs within each group.[22] On the other hand, since I treat all individuals within a group as similar individuals that share the same values of norms wedges, group market productivities, and group home productivities, it must also be that the categorization of the group is specific enough.[23] Of the variables defining each group, average schooling has undergone drastic increases in the sample period. I therefore adjust for the fact that the commonly completed levels of schooling differ by decade. I construct 5 or 6 schooling levels every decade, with at least 5% of the sample belonging to each level. Similarly, I construct family composition categories to have group sizes that are large enough and group categories specific enough. As the largest differences in home production duties relating to children occur for the first child and whether the child is young, these factors formed the basis of the categorization.[24]

Figure 5 plots the weighted average of group market wage and home productivity by sex and decade.[25] The group market productivity $\bar{w}$ is increasing in the average wage of the workers in that group as well as in the LFP rate of that group. The reason $\bar{w}$ increases in group LFP rate is that the group LFP rate encompasses the selection effect. The higher the LFP rate, the lower the average idiosyncratic market ability, as the set of workers are less selected on market ability. Then for the same empirically observed average wage of those who work in a group, less of it is accounted for by the average idiosyncratic market ability, so the higher the group market productivity must be. The group home productivity $\bar{h}$ is also increasing in the average wage of the workers in that group, but is decreasing in the labor force participation rate of that group.

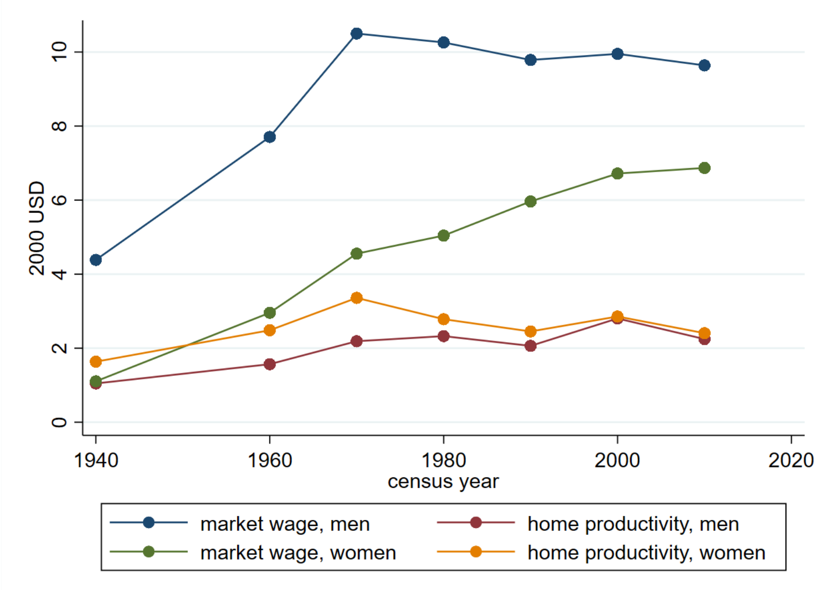
[22] For example, see equations (10) and (11), applied in steps 3 and 4 of section 4.2.

[23] Table A1 of the Appendix demonstrates that education is by far the most important predictor of attitudes relating to the role of married women, followed by survey year.

[24] The categorization of schooling levels and family composition are in Appendix tables A2 and A3.

[25] The weight equals the empirical probability of each group. For example, if the share of college-college couples with no child among the entire sample is high in 2010, the group market wage received by the wives of such couples get a greater weight in computing the average group market wage for women in 2010.

Figure 5: Weighted Average of Group Market Productivity ($\bar{w}$) and Home Productivity ($\bar{h}$) by Sex



Note: This figure plots the weighted average of the group components of market productivity and home productivity by sex and decade. These productivities for each group are inferred using the model structure as outlined in steps 3 and 4 of section 4.2.

How do $\bar{w}$ and $\bar{h}$ vary by education? As mentioned in section 3.7 on the intuition for aggregate productivity effects of decreases in norms wedges, whether an increase in education increases $\bar{w}$ or $\bar{h}$ by more matters for the educational choice in counterfactual scenarios. Specifically, if a young woman anticipates a higher likelihood of market work due to a fall in norms wedges, she will increase her education if education increases market productivity by more than home productivity. Table A4 of the Appendix confirms that for both sexes in every year, education increases market productivity by more than home productivity.[26]

[26] In fact, while market productivities significantly increase with education every decade, home productivities are either not affected or *decreasing* in education other than for women in 1940 and 1960. This finding contrasts with the literature on the positive returns to education on childcare (Leibowitz, 1974). It could be that $\bar{h}$ is underestimated, and more so at higher levels of education, in the later decades. $\bar{h}$ is inferred from singles' labor force participation behavior. If with greater marketization over time (e.g. Ngai and Petrongolo, 2017) it becomes more important for a household to have a wage income, then singles will be more likely to be in the labor force than married individuals as they do not have spouses that can bring in the wage income. Then using the inferred $\bar{h}$ from singles for the $\bar{h}$ for marrieds will underestimate the home productivity for marrieds. Moreover, this underestimation may be more pronounced at higher levels of education where a higher wage income is at stake in the singles' labor supply decision.

τ: gender norms wedges

I next calibrate the values of τ, the norms wedges on the market wage of married women. To illustrate the interpretation of τ through the lens of the model, $\tau = 0.6$ indicates that the worth of a \$10 market wage to a married woman is only \$4 when she is making her labor supply decision.

Though labeled as "norms wedges", τ is a catch-all term that encapsulates any reason that brings married women's LFP to diverge from single women's LFP besides the wage differentials. Therefore, it is helpful to think carefully about what factors do and do not go into τ. Because the effect of wage differentials are removed out, τ is *not* confounded by the fact that married women might have different market abilities than similar single women. On other other hand, τ does include many things that fall under the broad concept of homemaker norms for married women, either self-imposed or expected by others: a) the preference to conform with traditional gender norms as a wife, b) differential valuation of staying home for married women relative to single women,[27] c) differential non-wage treatment of married women relative to single women by firms, and d) men's preference to have more home-productive women as wives. All these factors support the labeling of τ as "norms wedges." One factor τ includes, which categorically does not fall under the broad idea of homemaker norms, is married women's LFP being affected by the fact that they operate within a two-person household where specialization is possible, whereas single women operate within a one-person household. I address this issue by calibrating also a model extension (Appendix section C.4) that directly incorporates this difference, discussed further below. Moreover, I show in Appendix section E.2 how τ is robust to the removal of two other factors unrelated to culture setting married women apart from single women: differential tax treatment by marital status and differential child-bearing prospects.

Figure 6 plots the histograms of τ, calibrated by group, for the years 1940 and 2000. The height of the bar for each group equals the group's empirical probability. It is very noticeable that the histogram of τ for 2000 sit to the left of that for 1940, signifying a decrease in the norms wedges.

I take the inferred group-specific τ as noisy estimates of norms wedges, as there are many selection effects unaccounted for in the model, such as the correlation between education choice and market abilities, or the correlation between taste for spousal type and market abilities. Therefore, I consider the weighted median of τ by decade.[28] Figure 7 plots these values in blue. τ generally decreases over time, except from 2000 to 2010. The mathematical

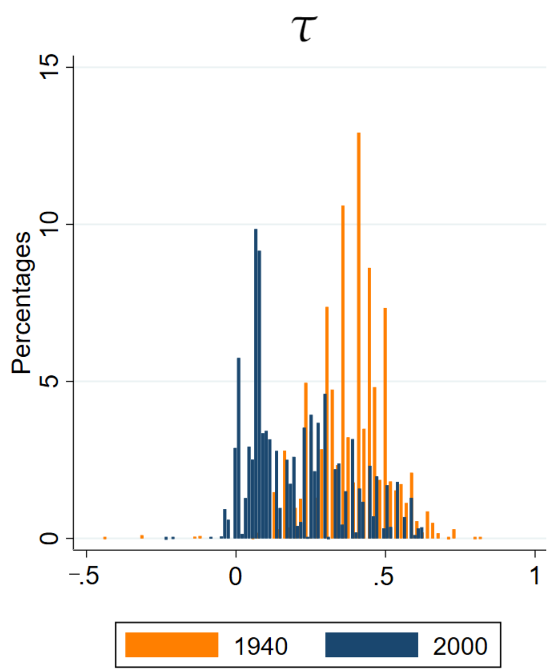
[27]See Appendix section C.3 for an explicit demonstration.

[28]The results are very similar to weighted average, unweighted average, or unweighted median values.

reason for the rise at the end is that while married women's wages have increased relative to single women's between the two decades, their LFP has not.[29,30] To reconcile these two observations, τ must increase.

The comparison of the τ's in the benchmark model to the counterpart τ values in the model extension with income effects (Appendix section C.4), in Figure 7, reveals more. In the extension, married women's labor supply decisions respond to their husbands' market earnings; (household) income effects decrease married women's incentives to bring additional dollars into the household. τ does not rise between 2000 and 2010 in the model extension, showing that it is the failure to capture spousal dependence that overestimates τ in 2010 in the benchmark model. Nonetheless, the benchmark model performs well up to 2000.

Figure 6: Histogram of Norms Wedges on Married Women's Market Wages (τ)

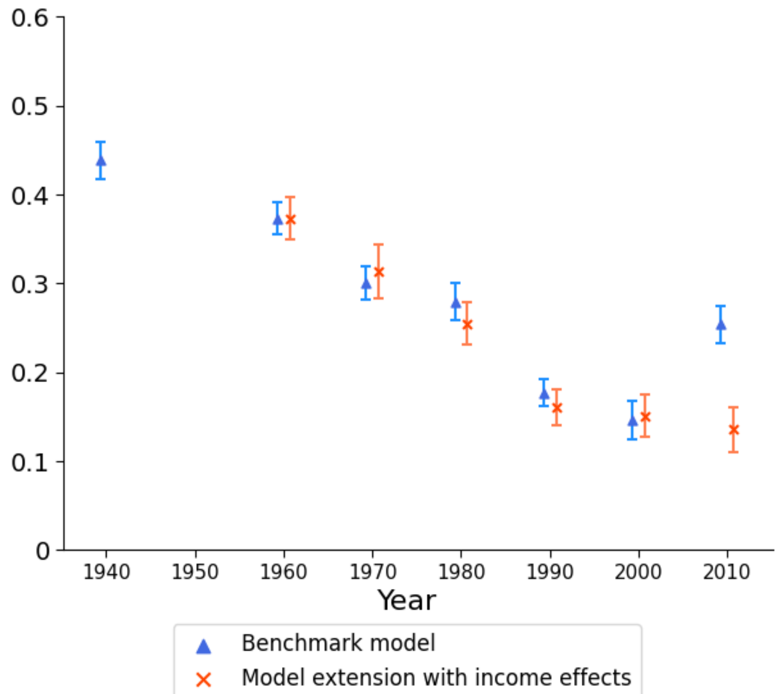


Note: This figure plots the histogram of τ, calibrated by group, for the years 1940 and 2000. The height of the histogram bars is scaled to percentages so that it indicates the empirical probability of the corresponding group in each year. The norm wedges for each group are inferred using the model structure as outlined in step 5 of section 4.2.

[29] Figure 1 illustrates that female labor force participation has plateaued since 1990.

[30] The rise persists even when I take into account the market work *hours* of married women relative to single women.

Figure 7: Evolution of Norms Wedges τ



Note: This figure plots the weighted median of τ by decade, in the benchmark model and in the model extension with income effects (Appendix section C.4). The weight equals the empirical probability of each group that τ is inferred for. The error bars indicate 95% confidence intervals based on bootstrapped standard errors with 100 replications. The τ in the model extension can only be estimated from 1960, because household income beyond one's own earnings (spousal earnings, welfare benefits, social security income, unearned income) are reported only from 1960.

To show that the values of norms wedges imputed are related to directly observable measures of conservativeness, I correlate state-level τ's with state-level attitudinal survey answers in table 2.[31,32] Two attitudinal measures are considered. The first is the fraction disapproving of married women working, from the single attitudinal survey question featured in Figure 3. The second is a composite attitudinal index that takes the weighted average by state of all the attitudinal survey questions plotted in Figure A1 of the Appendix. It is reassuring that the states with more conservative gender attitudes are also the ones with higher norms wedges τ.

[31] I redo the parameter inference procedure at the state level while pooling all the data across the years. As before, to ensure that each group is specific enough but also large enough, group categories are reformulated: defined by state, schooling pair, and age cohorts (ages 25-34, 35-44, and 45-54.) Then I take for each state either the weighted average or the weighted median of the norms wedges τ.

[32] The state-level attitudinal survey answers are the weighted average by state of individual survey answers taken across multiple periods.

Table 2: Correlation between State-Level Norms Wedge and Attitudinal Survey Answers

	Dependent variable τ	
	average	median
Regressed on:		
Fraction disapproving of	0.249**	0.282**
married women working	(2.21)	(2.21)
Regressed on:		
Composite attitudinal index	0.450***	0.439**
	(2.94)	(2.50)
N	51	51

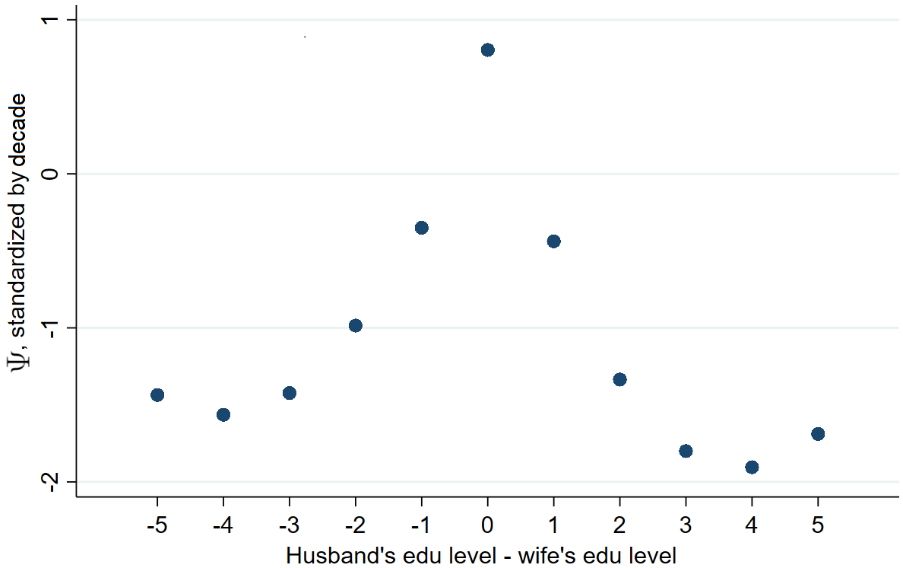
Note: t statistics using robust standard errors in parentheses, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$
Higher values of the composite attitudinal index correspond to more traditional attitudes. Each group gets weight equal to empirical probability in computing the average and the median. The positive correlation between τ and the attitudinal survey answers survives when the regressions are run at the state-year level with year fixed effects. Adding state fixed effects to this regression, however, renders the coefficients statistically insignificant, due to the fact that τ is a noisy measure of traditional gender norms.

Ψ: utility from marital bliss and intra-household resource allocation

Figure 8 plots the average Ψ for each value of the difference in the husband's and wife's education levels. As the values of Ψ differ by decade, driven by changing marriage patterns over time, Ψ are standardized by decade before averages are taken across the decades. The reason for Ψ at the spousal difference of 5 being larger than Ψ at the difference of 4 is that there are 6 schooling categories only in 2 decades, while the other 5 decades have 5 schooling categories. Overall, this plot can be taken as a single-peaked plot, peaking when the husband and wife share the same education level. Therefore, the calibrated Ψ values are congruous with the well-documented fact of assortative matching by education in the U.S. (Greenwood et al., 2016).

To show how the pattern of Ψ has changed *over time*, Figure A5 of the Appendix decomposes the content of Figure 8 by decade. The values in Figure A5 are *not* standardized by decade, and hence the fall in the values over time reflect declining attractiveness of marriage.

Figure 8: Ψ by Spousal Education Gap



Note: This figure plots the average Ψ for each value of the difference in the husband's and wife's education levels. As the values of Ψ differ by decade, driven by marriage patterns changing over time, Ψ are standardardized by decade before averages are taken across the decades.

c: utility cost of schooling

The last set of parameters to calibrate relates to the gender-specific cost of schooling, c_g^s. For example, the cost to females of acquiring the schooling level of "high school graduate" is inferred to be small if there are more female high school graduates in the data than is predicted by the expected utility from that level of schooling. The expected utility depends on the marriage market returns and the economic (i.e. wage and home productivity) returns.

The estimates of c, reported in Appendix Table A5, increase in the level of schooling for each year. Moreover, the costs fall over time in general, matching the empirical pattern of the rapid rise in average education levels. In addition, unlike how men's cost of acquiring the highest level of schooling does not show a clear pattern of decrease over time, women's does. This observation matches the stylized fact of women's overtaking of men in educational attainment in the U.S. For instance, the share of 25- to 34-year-old women with at least bachelor's degrees overtook that of men around 1995.

5 Counterfactual Exercises

In order to quantify the contribution to economic growth of changes in gender roles, I will consider how aggregate output Y changes if the norms wedges are the only parameters changing while all others are kept fixed. Moreover, I benchmark the effects of this main counterfactual on the effects of other counterfactuals described below.

5.1 Steps for conducting counterfactuals

I denote counterfactual values with underlines.

1. Compute $\underline{A}^{qr}$ and $\underline{P}_g^{qr}(\mathcal{K})$. They differ from the values at the status quo because optimal labor decisions would change under different gender norms. Note that $\hat{A}_g^s$ remains unchanged, as singles' labor decision is unaffected by norms wedges.
2. Compute counterfactual schooling probabilities $\frac{\underline{F}^r}{\sum_{r'=1}^{S}\underline{F}^{r'}}$ and $\frac{\underline{M}^q}{\sum_{q'=1}^{S}\underline{M}^{q'}}$, using $\underline{A}^{qr}$. Then compute $\underline{F}^r$ and $\underline{M}^q$ by assuming that the total population size remains constant.
3. Solve for the $(S\times S+2S)$ values of marriage matches $\underline{n}^{qr}$, from the $S\times S$ equations given by (19) as well as the $2S$ accounting identities given by (12) and (13).
4. Finally, compute $\underline{Y}^{mkt}$ and $\underline{Y}^{home}$

5.2 Counterfactual exercise results

Table 3 records the changes in various aggregate variables that would occur in 2000, the latest decade in which the benchmark model performs well, if gender norms had remained at the level of 1940. All other parameter values are held fixed at the 2000 level.[33]

I consider two adjustment margins: a) when only the labor supply choices are allowed to respond, in column (1), and b) when education, marriage, and labor supply choices are all allowed to respond, in column (2). Because not all variables change in column (1), column (1) clarifies which variables are directly affected by τ. When τ increases from 0.15 in 2000 to 0.44 in 1940, i.e. more traditional gender norms for married women, married women work 24.3% less in the labor market. As a result, the cumulative market output Y^{mkt} of married women falls by 12.5%. However, as fewer married women work in the market, married women's cumulative home production value Y^{home} increases, and so the total output

[33] As discussed in section 4.3, the calibrated norms wedges for *each group* are most likely noisy estimates. For this reason, I consider the counterfactual of every individual in 2000 being subject to the same, weighted-median norms wedge of 2000 at baseline, and being subject to the same, weighted-median norms wedge of 1940 in the counterfactual scenario.

$Y(= Y^{mkt} + Y^{home})$ falls by less, at 3.1%. The dissimilar effects on Y^{mkt} and Y highlight the importance of accounting for nonmarket output, which is almost always excluded from national accounts. It therefore hints that the output gains when women enter the labor market would be overstated in methods that only consider market output.

Table 3: Percent Changes in Various Aggregate Variables if Individuals of 2000 were Subject to the Norms Wedge of 1940

	Adjustment margins	
	Labor supply (1)	Education, marriage, & labor supply (2)
Education		
Women's average years of schooling	-	-2.1
Men's average years of schooling	-	-5.5
Selection into marriage		
Marriage rate	-	-33.4
Married women's average years of schooling	-	-8.1
Single women's average years of schooling	-	4.0
Married men's average years of schooling	-	-6.8
Single men's average years of schooling	-	-3.9
Labor Force Participation		
Married women's LFP	-24.3	-34.1
Single women's LFP	-	1.6
Married men's LFP	-	-2.8
Single men's LFP	-	-1.4
Output per head		
Married women's market output	-12.5	-28.2
Married women's total output	-3.1	-12.7
Married men's market output	-	-10.8
Married men's total output	-	-8.7
Aggregate market output	-3.6	-9.0
Aggregate market & home output	-1.0	-5.5
Within-household gender earnings gap		
Wife's share of household market income	-20.0	-25.8

Note: This table reports the percentage changes in various aggregate variables that occur when the individuals of 2000 are subject to the norms wedge of 1940, holding all other parameter values constant at the 2000 level. Column (1) holds the marriage match patterns and educational choices constant at the 2000 level and considers only changes to the married individuals' labor supply decisions. Column (2) additionally allows the (forward-looking) marriage matching and educational choices to change in accordance with the new expected utilities arising from the altered labor supply behavior.

In column (2), the direct effects of higher τ trickle down to indirectly affect education and marriage match choices, too. As norms wedges are modeled as costs to marriage, a higher τ renders marriage less attractive, yielding a fall in the marriage rate of 33.4%. As the norms wedge is more costly for women with higher market wages, higher-educated women are the first to drop out of marriage. Due to assortative marriage matching by education, the fall in marriage is more pronounced also for higher educated *men*. As the set of married women is now more negatively selected on education, married women's LFP rate and cumulative market output fall by *more* in column (2) than in column (1). Married women's cumulative total output falls by 12.7%. Furthermore, the education of both men and women fall on average. The fall, however, is less pronounced for women. While men's education choices are only affected by lower marriage market returns, the lower marriage market returns for women are countered by higher labor market returns. Highly educated women are less likely to get married, no longer subject to norms wedges and more likely to be working. In total, aggregate market output falls by 9.0% and total output by 5.5%.

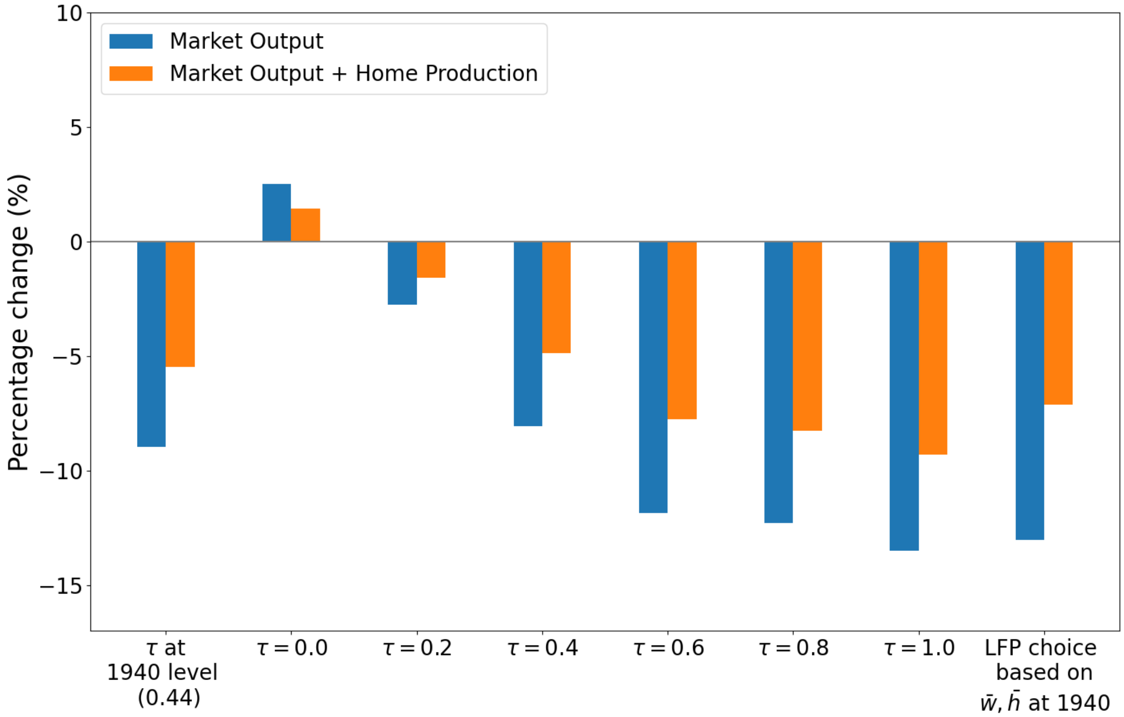
In summary, the fall in gender norms wedges over 1940 through 2000 partially accounts for various stylized facts documented in the U.S.: a) rise in married female LFP, b) rise in wife's share of household market income, c) growing educational attainment of both men and women, and d) increasingly positive selection of men and women into marriage by education (Bar et al., 2018; Juhn and McCue, 2017; Case and Deaton, 2017).

Is the effect of a 9.0% fall in aggregate market output and a 5.5% fall in aggregate total output small or large? The output effects might be viewed as large, as the norms wedge parameters capture quite a narrow concept of gender norms relating to the distinction between married and single women who are similar. On the other hand, it might be viewed as small, relative to the output growth that has occurred over 1940-2000. To better benchmark the size of the output effects, I conduct additional counterfactuals.

Additional Counterfactuals Figure 9 compares the market and total (market and non-market) output effects of various counterfactual scenarios, where the baseline year is set at 2000. The other counterfactual scenarios explore the effects of τ changing from 0 to 1, and when labor supply choices of females are made based on $\bar{w}$ and $\bar{h}$ of 1940. All the three margins of education, marriage, and labor supply are allowed to adjust.

As τ increases, the negative output effects becomes increasingly pronounced. When married women are completely prohibited from working (τ=1), market output falls by 13.5% and total output by 9.3%. This counterfactual highlights the potential loss in aggregate output incurred in the countries with the most conservative attitudes towards married women working.

Figure 9: Market and Total Output Effects of Various Counterfactual Scenarios



Note: This figure compares the market and total (market and nonmarket) output effects of various counterfactual scenarios, where the baseline year is set at 2000, with τ at 2000 being 0.15.

I focus on the last counterfactual, in particular, to benchmark the effects of the main counterfactual. As illustrated in Figure 5, the market and home productivities have undergone substantial changes over time, driving the economic growth over 1940-2000. In fact, $\bar{w}$ is lower than $\bar{h}$ for women in 1940, whereas $\bar{w}$ is *well over double* of $\bar{h}$ in 2000. Not to include the direct effects of productivity changes on output, the counterfactual is only about letting female labor supply choices be determined based on $\bar{w}$ and $\bar{h}$ at 1940 levels, while wages and home productivities remain at the 2000 level. The labor supply choice changes enormously, with married women's LFP rate falling by a staggering 70%, about double the drop in the main counterfactual. Market output consequently falls by 13.0% and total output falls by 7.1%. Therefore, the main counterfactual's output effects amount to nearly 80% of the output effects of the last counterfactual. In this sense, the effect of the norms wedge is sizable.

6 Reduced-Form Exercise

The counterfactual results in the previous section depend on the model structure. The extended discussion on the calibrated parameter values in section 4.3 describes how they match similar estimates in the literature, related measures from external data sources, or well-documented stylized facts. Yet, to provide *further* evidence in support of the model, I perform a reduced-form exercise. The exercise also allows me to add an extension to the model where *economywide* gender norms respond to *economywide* past female labor force participation, and using this relationship, to conduct dynamic counterfactuals.[34]

6.1 Model validation

For lack of a direct test of model predictions when norms wedges fall, I explore the effects of a shock that *indirectly* affects norms and check that other variables change in the expected direction. To clarify, the purpose of this exercise is not to demonstrate that the effect of this shock is entirely explained by norms, but rather that the shock triggers changes that are consistent with model predictions when norms change.

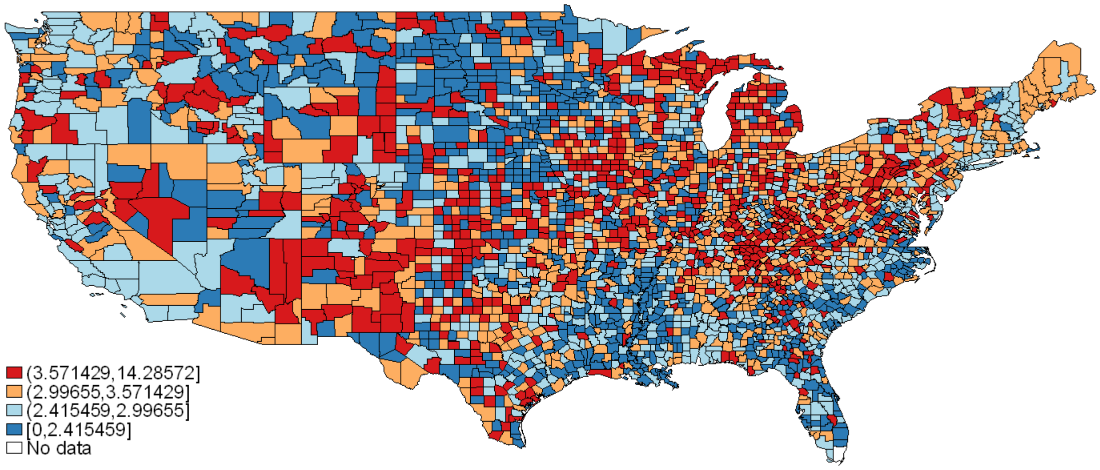
Inspired by Fernández et al.(2004),[35] I take WW2 draftee casualties as a temporary positive shock to female labor force participation that propagates over the long term through weaker gender norms. Underlying this story is the idea of cultural transmission through exposure (Bisin and Verdier, 2000) or social learning (Fogli and Veldkamp, 2011; Fernández, 2013).

For the reduced-form exercise, I match the U.S. decennial census, by county, with the WW2 military casualty records from Ferrara (2021). As a result, every county is characterized by the casualty rates among draftees, as illustrated in Figure 10. While earlier studies on the effects of WW2 utilized WW2 mobilization rates by state (e.g. Acemoglu, Autor, and Lyle, 2004; Fernández, Fogli, and Olivetti, 2004), newly digitized data from the National Archives and Records Administration enables the use of *county*-level variation. Moreover, there are two advantages to using casualty rates as opposed to mobilization rates. First, although most women who engaged in wartime work left the labor force upon demobilization (Goldin, 1991), casualties last. Second, casualty rates among draftees are likely to be more random than mobilization rates.

[34]This model extension is compatible with how I identified norms wedges previously, as long as each individual takes norms as given and do not internalize the effect of their labor supply choice on the norms of future generations.

[35]They argue that WW2 mobilization weakened traditional gender norms over the long run, as sons of women who worked during the war grew up to be more accepting of working wives.

Figure 10: Map of County-Level Draftee Casualty Rates



Note: This figure color-codes each county into quartiles of draftee casualty rates. From the highest to the lowest quartile, the colors are red, orange, light blue, and blue.

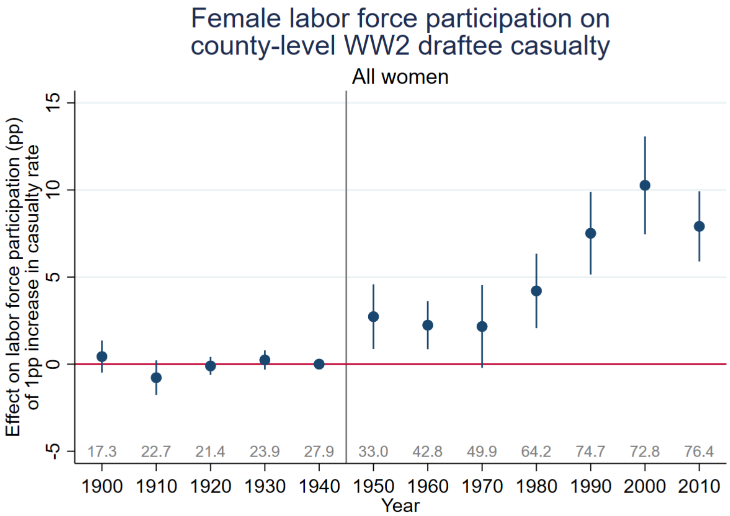
The baseline estimation strategy for the effect of draftee casualties is difference-in-differences with continuous treatment. Hence, I estimate, for individual i in county c at decade t,

$$Y_{ict} = \alpha_c + \lambda_t + \sum_{t \neq 1940} \beta_t \times casualty_c + X_{ict}\gamma + \varepsilon_{ict} \tag{20}$$

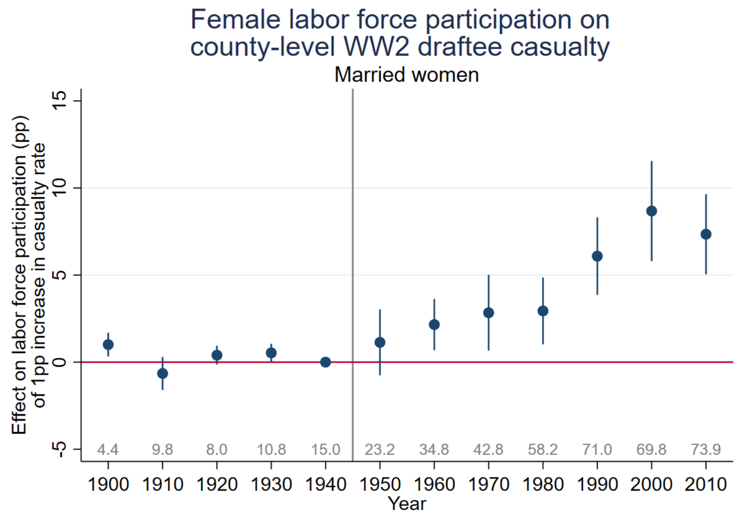
where Y represents various outcome variables, *casualty* is the county-level draftee casualty rate – the death rate among all soldiers drafted from a county – and X captures predetermined individual characteristics, namely dummies for age and race, added for greater precision. In other words, I study the effect of the casualty rate in each decade t relative to 1940, the last decade before the influence of WW2 reached the U.S. With parallel trends, it must be that $\beta_t = 0$ for all $t < 1940$. I later check that the results according to the specification in (20) are robust to a) comparing above-median- to below-median-casualty counties in a standard binary difference-in-differences framework, b) controlling for 1940 county characteristics that predict casualty rates, interacted with decade dummies, in order to address the nonrandomness of casualty rates, and c) applying the synthetic difference-in-differences methodology (Arkhangelsky et al., 2021) to further allay concerns over *level* differences in the pre-WW2 period affecting the future trajectory of various outcome variables.

The main results are depicted in Figure 11. Plot (A) shows that a 1 percentage point increase in draftee casualty rate induces a 2.5 percentage point increase in female labor force participation rate in 1950. When I dissect the source of this spike, it comes from widows and single women living with their parents increasing their labor force participation, consistent with firms demanding more female labor with lower male labor supply, and new widows

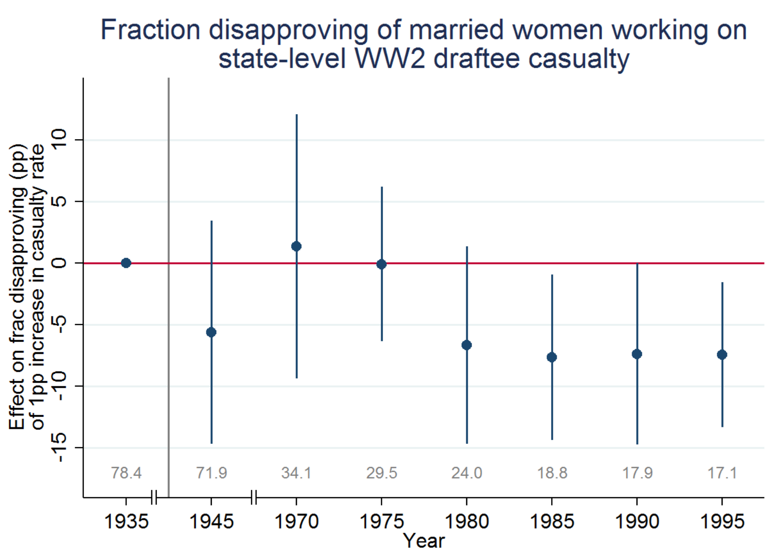
Figure 11: The Effect of WW2 Draftee Casualty Rates on Various Outcomes



(A)



(B)



(C)

Note: This figure plots the difference-in-differences coefficients from estimating equation (20) for various outcome variables. Plot (C) uses state-level draftee casualty rates, because state is the finest geographic variable available in attitudes data prior to WW2. In Plot (C), 1945 is counted as post-WW2, since the survey was taken in November, after the official end of WW2 in September.

increasing their labor supply as a direct consequence of the casualties. The effect of casualties on female labor force participation in 1960 is still positive but a little smaller than in 1950, and then displays a gradual increase over the next decades. Plot (B) shows that for the most part the gradual increase is driven by married women. At the same time, plot (C) shows that gender attitudes become gradually less traditional with higher casualties.[36] Put together, the three panels are consistent with a story of a one-off rise in female labor force participation propagating over the long term through less traditional gender norms that primarily affect married women's labor force participation.

Appendix figure A6 portrays the reduced-form results for other variables that buttress this story as well as the model structure. Single women's labor force participation in plot (B) does not mimic the strong, gradual rise observed for married women, reproduced in plot (A). This contrast supports the model assumption where only married individuals' labor force participation decisions are affected by the norms in marriage. Men's employment, as shown in plot (C) does not exhibit a systematic change over time, indicating that the gender norm change is associated with a change in women's behavior, mostly. Plot (D) shows that the wife's share of the couple's wage income is also increasing over time, even as real hourly wage for working women is decreasing over time, shown in plot (E). Although not included, the wife's share of the couple's total market hours worked also increases over time, gradually and continually. Furthermore, while wages are equilibrium prices jointly determined by supply and demand, the decrease in female wages points more towards a rise in female labor supply than female labor demand driving the rise in female labor force participation. The effect on female wages thus supports the model assumption of a decrease in norms wedge affecting the labor *supply* decisions of married women. In addition, the decrease in female wages is consistent with the model assumption around selection into the labor force, i.e. as more women work, working women are less positively selected. The observation that there are long-term differences in female wages between high- and low-casualty counties indicates that there are labor market frictions precluding the equality of wages across space. In fact, there is no consistent trend of individuals moving to either high- or low-casualty areas to wash out the effects of WW2 casualties, which would have shown up in people migrating to different states from their birth states in plot (F).

In terms of marriage and education, plot (G) depicts a rise in ever-marriage rates and plot (H) a rise in the education of women overall. Plot (H) uses whether one graduated high school or more as the measure of education, as higher levels of education are very rare to

[36]There are a few caveats for plot (C). As the finest geographic variable in the attitudinal survey data is state, the difference-in-differences analysis is performed using state-level draftee casualties. Moreover, the surveys are grouped into five-year intervals. 1945 is counted as post-WW2, since the survey was taken in November, after the official end of WW2 in September. The statistically insignificant drop in 1945 appears to be out of trend, but it is the date in which the sample is by far the smallest; the sample is 1,365 in 1945, while the other dates are based on around 3,000-6,000 observations. 1945 is also the only date in which no survey weights are available. Overall, I take the coefficient plot of plot (C) to indicate that the attitude data is quite noisy, and that the attitudes getting less traditional becomes detectable (statistically significant) from 1985.

find in 1940, the pre-WW2 benchmark decade. The effect on marriage is consistent with the model assumption where norms wedges are modeled as costs to marriage, and as the norms wedge falls over time, more people engage in the tradition of marriage. Also, the effect on education supports the model prediction of greater female education due to the greater likelihood of marriage and market work. Lastly, the model assumes that the norms wedges impose higher costs on women with higher market ability. This assumption leads to the model prediction of increasingly positive selection into marriage by education of women, as the norms wedge falls. The rise in education of *married* women, in plot (I), but not for single women, in plot (J) supports this prediction.

All in all, it is difficult to reconcile how women are getting married and educated more, married women but not single women are working in the market more, and female wages falling, without a story of changing gender norms. Gender attitudes indeed become less traditional over time in the data. Surely, WW2 casualties can have alternative effects. For instance, the fall in sex ratio can increase husbands' bargaining power. Yet in that case, married women would *not* increase market work, since attitudinal surveys indicate that men hold more traditional views on married women working than women. As another example, casualties might somehow change the industrial structure into one that better enables women to combine work and marriage, such that women with higher market ability get married more. However, while this might explain the rise in married women's market work, the rise in marriage, and the rise in female education, it goes against falling female wages. To generate the sizable rise in married women's work *soley* from a higher market talent of married women, married women's market talent must rise by a great amount, in which case it is unlikely to see a fall in female wages.

Robustness I firstly check that the effect of WW2 draftee casualties survive a binary difference-in-differences specification. Column (1) of Appendix table A6 reports the effects of casualties on female labor force participation, pictured in plot (B) of figure 11. Column (2), which reports the binary specification results, are very similar to column (1).

Secondly, I control for 1940 county characteristics that predict casualty rates, interacted with decade dummies, to address the nonrandomness of casualty rates. Indeed, the casualty rates are not completely random. Figure 10 shows spatial clustering in the casualty rates. During WW2, drafted soldiers were assembled at *state* base camps, and casualties were dictated by outcomes of specific battles, so nearby counties experienced similar casualty rates. Moreover, blacks were killed at a lower rate since they were mainly employed in comparatively safer support and supply activities due to racist attitudes that saw them unfit for fighting (Lee, 1965). Appendix table A7 shows that casualty rates were higher in counties with a higher share of whites, a lower share of working-age women, a higher urban resident share, a lower male education, and a lower share of men in agriculture. I therefore control for the effects of these 1940 county characteristics over time in column (3) of Appendix table A6. Although the coefficient sizes get smaller, the pattern of a jump up in 1950 followed by a gradual rise in female labor force participation remains.

Lastly, I apply the synthetic difference-in-differences methodology of Arkhangelsky et al. (2021), which re-weights and matches pre-exposure trends to weaken the reliance on the parallel trends assumption. Because of the need to divide counties into treatment and control groups, I can only employ a binary specification. Column (4) of Appendix table A6 shows that the synthetic difference-in-differences coefficients also depict the same patterns.

6.2 Dynamic counterfactuals

The reduced-form results based on WW2 casualties are useful for validating the assumptions of the structural model. Not only that, but those results also allow me to consider a dynamic extension to the model. Up to now, I have assumed that the model is in equilibrium each decade, with no dynamic element linking any two decades in the model. Yet, it is unlikely that gender norms evolve entirely exogenously. In fact, the WW2 reduce form results are congruent with temporarily higher female labor force participation inducing a gradual fall in the norms wedge.

I estimate how the norms wedge responds to past female labor force participation, using the reduce form results in plot (A) of figure 11. To this end, I first impose a functional form on how *economywide* norms wedge in decade t responds to *economywide* female labor force participation in decade $t-1$:

$$\begin{aligned} \underline{\Delta}\tau_t &= f(\underline{\Delta}FLFP_{t-1}, FLFP_{t-1}) + \nu_t \\ &\approx \alpha_0 + \alpha_1 \underline{\Delta}FLFP_{t-1} + \alpha_2 FLFP_{t-1} + \nu_t \end{aligned} \tag{21}$$

where $\underline{\Delta}$ denotes the gap between treatment and control counties. I also need two additional assumptions on how the long-term effects rise about: a) WW2 draftee casualties affect female labor force participation in 1950 and nothing else, and b) the effect only propagates through a change in the norms wedge. These assumptions allow me to take treatment and control counties as two identical worlds whose female labor force participation started diverging as a result of a temporary exogenous shock in 1950 and continued to diverge from 1960 as a result of norms evolving according to (21). I estimate the relationship in (21) that would generate the pattern of difference-in-differences coefficients in plot (A):

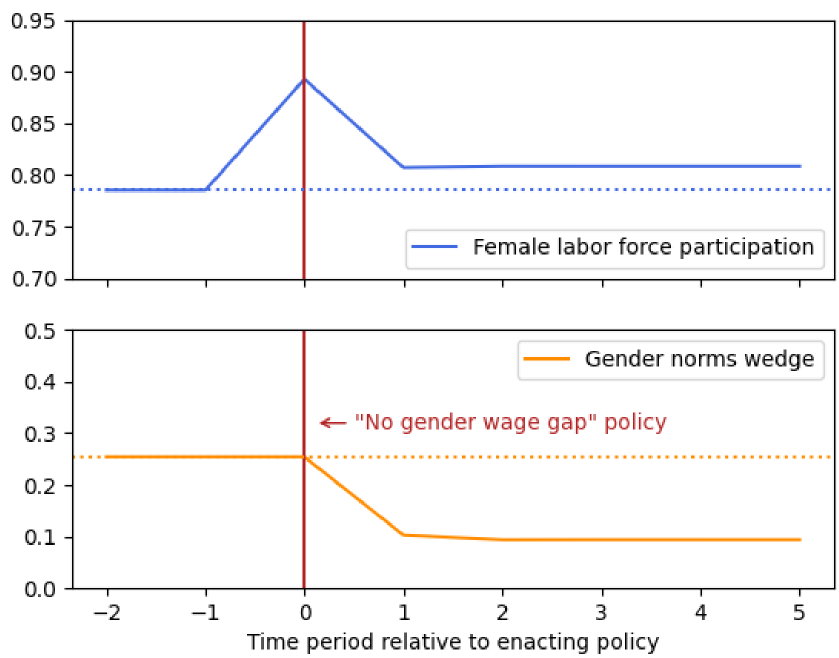
$$\min_{\alpha_0,\alpha_1,\alpha_2} \sum_{t=1960}^{2010} (\text{DID coeff for FLFP}_t - \text{change in FLFP}_t \text{ in model due to } \underline{\Delta}\tau_t)^2$$

The result is $\hat{\alpha}_0 = -0.119$, $\hat{\alpha}_1 = -0.468$, $\hat{\alpha}_2 = 0.359$.

Armed with this estimated relationship, I can further conduct *dynamic* counterfactuals. As opposed to the "static" counterfactuals on the effect of a shock on the model equilibrium in a given decade in section 5, I explore how a shock affects the model equilibrium *over time*.

The counterfactual I ask is, what would happen in 2010 if women were paid male wages in a one-off fashion? The counterfactual abstracts from labor demand being affected, as a consequence of the model assumption of firms producing under a linear production function. Forcing firms to pay women male wages, all the more without changing employment, is a far-fetched idea. Yet this thought experiment illustrates how a one-off policy can move an economy to a different equilibrium. Figure 12 shows that while keeping all other parameters fixed at the 2010 level, paying women male wages for one period induces a contemporaneous spike in female labor force participation, which then induces the norms wedge to fall a decade

Figure 12: The Effect of Paying Women Male Wages, One-Off, in 2010



Note: The figure plots the dynamic effect of a temporary "big-push" policy that greatly increases female participation. One decade after the policy, female labor force participation is higher than the pre-policy equilibrium level because the gender norms wedge have declined in response to the rise in female labor force participation in the previous decade. The new equilibrium the economy stabilizes to features a higher female labor force participation rate and a lower gender norms wedge.

later. The female labor force participation that decade is lower than the decade of the policy as the direct effect of the policy is gone, but it is higher than the baseline due to the lower norms wedge. Consequently, in the second decade after the policy, the norms wedge falls even more. The process continues, and from three decades post-policy, the economy stabilizes at a new equilibrium with higher female labor force participation and lower norms wedge than the baseline.

7 Concluding Remarks

In this paper, I measure and study the effects of gender roles associated with marriage on aggregate output, using historical data from the United States. Gender norms became less of a constraint on married women's labor supply choices over the last 70 years. Through the direct effects of improved alignment between labor supply choice and productivity maximization, and the indirect effects of higher human capital accumulation, less conservative gender norms boost aggregate market and total output.

Cultural costs on output are sizable. Operating under 1940 norms wedges in the current day drags aggregate output down by 5.5%. This amounts to nearly 80% of the output cost of labor force participation being based on the *economic* forces – wages and home productivities – of 1940. Then, is there anything to be done about norms wedges? Besides activist movements aimed at transforming mindsets, the paper suggests that a big-push labor policy can play a part. A one-off policy triggering a large rise in female labor force participation can bring an economy to a new equilibrium with higher female labor force participation and lower norms wedges permanently.

The tractability of the model and straightforward parameter inference procedure lend great practical advantages. For instance, Lee (2021) takes the model to 17 other countries and augments it with a male-female wedge, which is compared to the married women-single women wedge central to the current paper. On the other hand, there are other important dimensions that the model does not account for, such as occupations (Hsieh et al., 2019), work flexibility (Goldin, 2014), divorce (Fernández and Wong, 2014; Greenwood et al., 2016), and leisure (Aguiar and Hurst, 2007b). Building these factors into the model will allow a richer understanding of the effects of gender norms in marriage. I leave these for future research.

Appendix A Tables

Table A1: Variation in Attitudes by Individual Characteristics

	(1) Average	(2) F-statistic	(3) Shapley decomposition (%)
Year			
1930-1939	0.77		
1940-1949	0.71		
1970-1979	0.26	10.3	19.8
1980-1989	0.16		
1990-1999	0.14		
Marital status			
Married	0.19		
Widowed	0.28		
Divorced	0.18	1.0	2.3
Separated	0.19		
Never married	0.15		
Sex			
Male	0.20	14.2	1.1
Female	0.18		
Race			
White	0.28		
Black	0.25	14.1	3.2
Other	0.21		
Education			
Middle school graduate or lower	0.41		
High school drop-out	0.29		
High school graduate	0.20	153.6	62.6
College drop-out	0.14		
College graduate or higher	0.09		
Age			
20-29	0.23		
30-39	0.26	2.2	11.1
40-49	0.29		
50-59	0.33		

Note: This table reports by how much various individual characteristics account for the variation in *att*, the indicator variable for an individual's disapproval of a married woman working in the labor market, in the Gallup Polls and the General Social Survey. The specific attitudinal survey question is in Figure 3. Column (1) reports the weighted average disapproval rate by category of each variable, to show the variation across the categories. Columns (2) and (3) are based on the regression of *att* on the dummies for the categories of each variable, on the sample from the General Social Survey. Column (2) reports the F-statistic for the joint significance of the dummies, while column (3) reports the Shorrocks-Shapely decomposition denoting the relative contribution of each variable to the R squared.

Table A2: Years of Completed Schooling by Schooling Category by Year

Schooling level	Year						
	1940	1960	1970	1980	1990	2000	2010
1	[0,7]	[0,7]	[0,8]	[0,9]	[0,11]	[0,11]	[0,11]
2	8	8	[9,11]	[10,11]	12	12	12
3	[9,11]	[9,11]	12	12	[13,15]	[13,15]	[13,15]
4	12	12	[13,15]	[13,15]	16	16	16
5	$[13,\infty)$	[13,15]	$[16,\infty)$	16	$[17,\infty)$	$[17,\infty)$	$[17,\infty)$
6		$[16,\infty)$		$[17,\infty)$			

Note: Years of completed schooling are integers in each interval. For example, individuals with 0,1,...,7 years of completed schooling fall under schooling level 1 in 1940.

Table A3: Description of Family Composition Categories

Family composition categories	Description
1	No child
2	1 child, aged 6-18
3	1 child, aged 0-5
4	2 or more children, all aged 6-18
5	2 or more children, at least one aged 0-5

Note: All children are one's own children living in the same household.

Table A4: Coefficients from Regressing $\bar{w}$ or $\bar{h}$ on Schooling Level, by Sex and Year

		1940	1960	1970	1980	1990	2000	2010
Male	Market productivity, $\bar{w}$	0.61***	0.84***	1.53***	1.12***	1.91***	2.48***	2.60***
		(13.98)	(16.22)	(11.97)	(10.56)	(16.75)	(16.74)	(15.40)
	Home productivity, $\bar{h}$	0.021	-0.068**	-0.17***	-0.13***	-0.064	0.029	0.016
		(0.96)	(-3.16)	(-3.80)	(-5.26)	(-1.76)	(0.72)	(0.39)
Female	Market productivity, $\bar{w}$	0.25***	0.46***	1.03***	0.98***	1.71***	1.90***	2.06***
		(5.29)	(6.17)	(7.34)	(13.19)	(20.76)	(30.45)	(38.36)
	Home productivity, $\bar{h}$	0.20***	0.11**	0.085	-0.25***	-0.21***	-0.069*	0.029
		(8.52)	(2.90)	(0.90)	(-7.19)	(-6.02)	(-2.56)	(1.04)
Number of groups		210	210	150	210	150	150	150

Note: t statistics in parentheses, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

This table reports the coefficients on schooling level in the regression of either market wage or home productivity on schooling level. In this regression, each observation corresponds to each group, weighted by its empirical probability. As described in Appendix table A2, schooling level is a categorical variable with higher values representing greater completed years of schooling. For the sake of simplicitiy in showing how $\bar{w}$ and $\bar{h}$ vary by education, I treat schooling level as a continuous variable in these regressions. Since the schooling level formulations change over time, the coefficients are not directly comparable across the years.

Table A5: Estimates of the Costs of Schooling c by Sex and Decade

		Schooling level					
		1	2	3	4	5	6
1940	Male	-1.83	1.41	0.23	0.58	2.69	
	Female	-1.62	0.40	-0.14	0.26	2.32	
1960	Male	-1.37	-0.43	-0.61	-0.65	0.80	1.74
	Female	-1.25	-0.73	-1.03	-0.84	0.33	2.02
1970	Male	-1.17	-0.66	-0.97	0.34	1.40	
	Female	-1.07	-0.96	-1.22	-0.39	1.44	
1980	Male	-1.03	0.04	-1.77	-1.16	0.21	0.23
	Female	-1.59	-2.03	-1.31	-1.49	-0.81	-0.11
1990	Male	-1.96	-2.32	-1.68	-0.14	1.61	
	Female	-3.48	-2.50	-2.11	-0.42	0.39	
2000	Male	-2.60	-3.24	-1.81	-0.30	1.78	
	Female	-2.94	-2.74	-2.00	-0.69	0.12	
2010	Male	-2.71	-3.23	-1.93	-0.27	2.03	
	Female	-3.01	-2.58	-2.15	-0.51	0.42	

Note: This table reports the estimates of c_g^s, the utility cost of schooling by sex and decade. The schooling categories for each decade are described in Appendix Table A2.

Table A6: The Effect of WW2 Draftee Casualties on Female Labor Force Participation (percentage points)

	DID		DID, 1940 county controls	Synthetic DID
	continuous	binary	continuous	binary
	(1)	(2)	(3)	(4)
1900	0.43	0.95	-0.28	-
	(0.47)	(0.68)	(0.30)	
1910	-0.78	0.03	-0.16	-
	(0.51)	(0.68)	(0.32)	
1920	-0.10	0.11	0.04	-
	(0.26)	(0.38)	(0.28)	
1930	0.24	0.37	0.17	-
	(0.28)	(0.37)	(0.27)	
1950	2.72***	2.42***	1.74*	2.44***
	(0.95)	(0.91)	(0.91)	(0.81)
1960	2.24***	2.37***	0.53	1.29**
	(0.70)	(0.68)	(0.70)	(0.66)
1970	2.16*	2.58**	-0.47	0.49
	(1.21)	(1.14)	(1.44)	(1.24)
1980	4.20***	3.95***	1.26	2.22***
	(1.09)	(0.86)	(0.98)	(0.77)
1990	7.48***	6.50***	3.37***	5.42***
	(1.20)	(0.95)	(1.07)	(0.99)
2000	10.25***	8.99***	5.02***	8.02***
	(1.43)	(1.14)	(1.27)	(1.23)
2010	7.89***	7.38***	3.18***	5.76***
	(1.02)	(0.89)	(0.91)	(0.97)
N	2,508,006	2,508,006	2,418,946	2,102,295

Note: Standard errors (col (1)-(3): clustered by county, col (4): jackknife) in parentheses, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. This table reports the decade-specific difference-in-differences coefficients of female labor force participation on WW2 draftee casualties by county. Draftee casualties are formulated either as the raw continuous variable or as a binary variable equal to 1 if casualty rates are above the median. Columns (1) and (2) report the result of estimating equation (20). Column (3) adds as controls 1940 county characteristics that predict casualty rates, interacted with decade dummies, to address the nonrandomness of casualty rates. Column (4) reports the coefficients from synthetic difference-in-differences (Arkhangelsky et al., 2021) to weaken the reliance on parallel trends assumptions.

Table A7: WW2 Casualties and County Characteristics in 1940

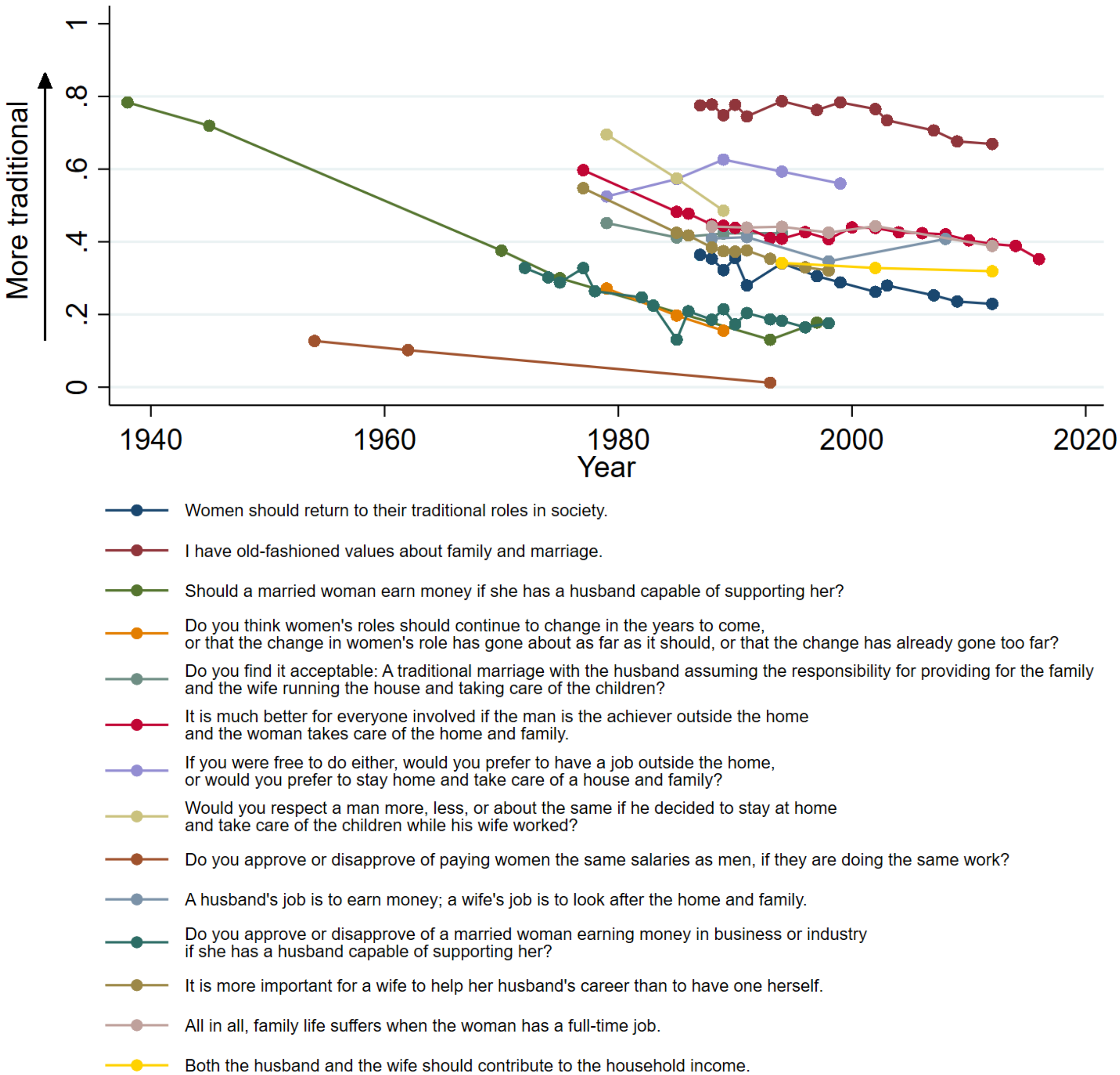
	Dependent variable: WW2 casualty rate among draftees, county-level			
	(1)	(2)	(3)	(4)
Share white	0.38***	0.32***	0.31***	0.31***
	(13.30)	(9.89)	(9.91)	(9.15)
Share aged 25-54 among women	-0.22***	-0.16***	-0.15***	-0.12***
	(-6.99)	(-3.75)	(-2.85)	(-2.71)
Share city resident	0.22***	0.23***	0.23***	0.18***
	(4.01)	(5.42)	(5.62)	(3.45)
Male avg. schooling	-0.13***	-0.09*	-0.14**	-0.14**
	(-3.07)	(-1.77)	(-2.32)	(-2.41)
Share in agriculture among men	-0.11***	-0.02	-0.03	-0.06
	(-3.88)	(-0.45)	(-0.77)	(-1.25)
Share married			-0.01	-0.01
			(-0.24)	(-0.17)
Female avg. schooling			0.07	0.05
			(1.02)	(0.92)
Avg. no. children in household			0.02	-0.02
			(0.51)	(-0.42)
Female labor force participation			-0.01	-0.02
			(-0.30)	(-0.57)
Additional controls	No	No	No	Yes
State dummies	No	Yes	Yes	Yes
No. counties	2409	2409	2409	2409
R^2	0.226	0.340	0.341	0.345

Note: t statistics (col (1): robust, col (2)-(4): clustered by state) in parentheses, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

This table reports the OLS coefficients from regressing county-level WW2 draftee casualty rates on various pre-war county characteristics. The sample excludes outliers, defined as counties having casualty rates or draft rates strictly outside the 2.5^{th} and 97.5^{th} percentiles. Additional controls are county population, share having non-wage income over \$50, share aged 25-54 among men, and share in agriculture among women. Average schooling, share in agriculture, share married, average number of children in household, and female labor force participation are computed among 25- to 54-year-olds.

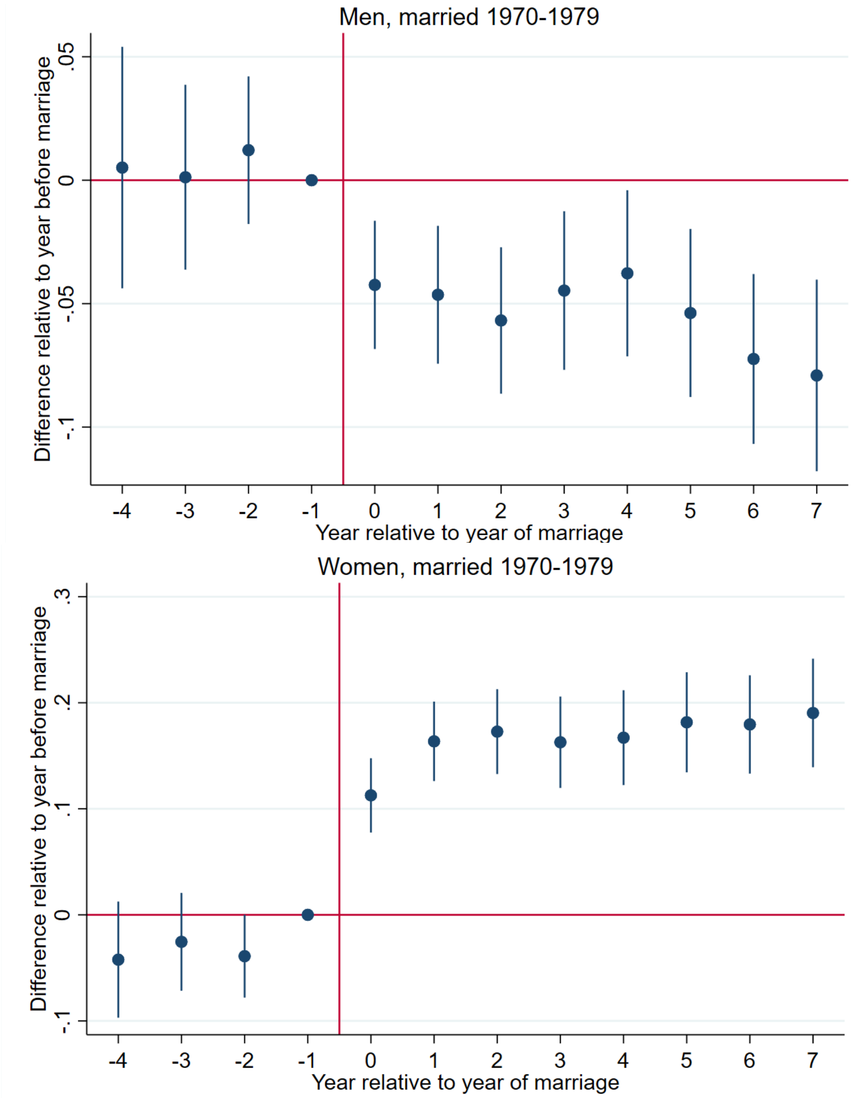
Appendix B Figures

Figure A1: Trends of Answers to Various Attitudinal Survey Questions Relating to Gender Roles within Marriage



Note: The attitudinal survey data originates from the ROPER polls database. I put together the surveys that come up from searches using the keywords "women" and "gender," and select the questions that relate to gender roles within marriage, and among them, the ones asked repeatedly over multiple periods. All samples are nationally representative adults or nationally representative men and women separately. All survey answers have been reconfigured to lie between 0 and 1, with 1 corresponding to the most traditional answer and 0 the least. This figure plots the weighted averages of the reconfigured answers by year.

Figure A2: Housework's Share of Housework and Market Hours, Among Couples Whose First Child is Born $\geq$ 4 years After Marriage

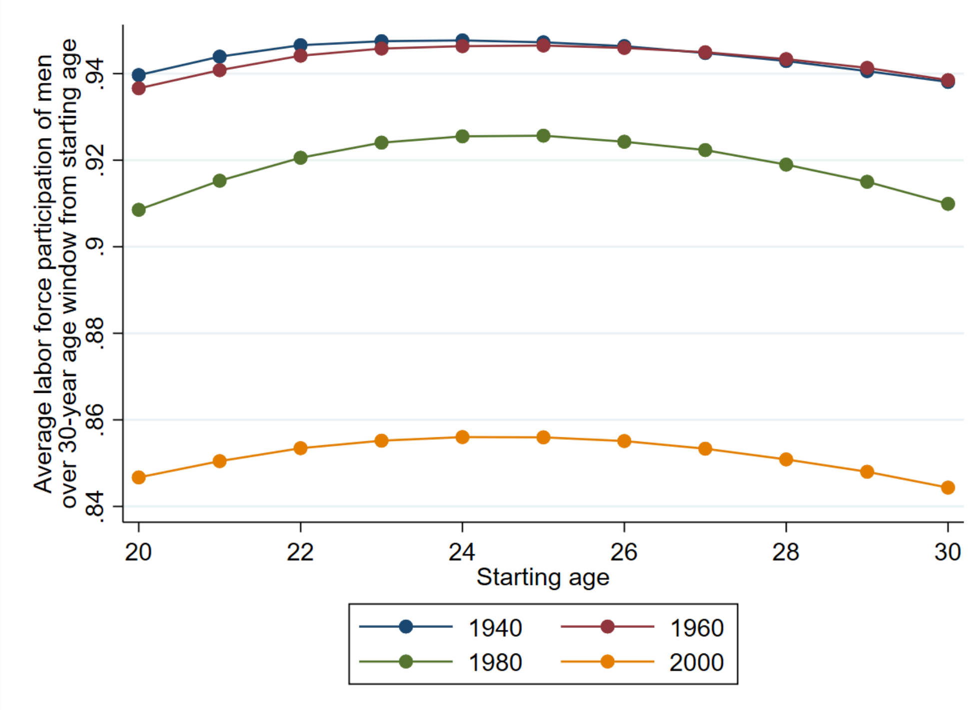


Note: This figure plots the event-time coefficients (α_j^g) of the regression

$$housework_{ist}^g = \sum_{j \neq -1} \alpha_j^g \cdot \mathbb{1}(j = t) + \sum_k \beta_k^g \cdot \mathbb{1}(k = age_{is}) + \sum_y \gamma_y^g \cdot \mathbb{1}(y = s) + \nu_{ist}^g$$

where $housework_{ist}^g$ denotes the housework's share of housework and market work hours of individual i of gender g in year s at event time t. The red vertical line plots the timing of marriage. Individuals are unmarried household heads without any live-in partners in the years to the left of the red line, and they are married with live-in spouses in the years to the right of the red line.

Figure A3: Average Labor Force Participation of Men over 30-year Window from Starting Age

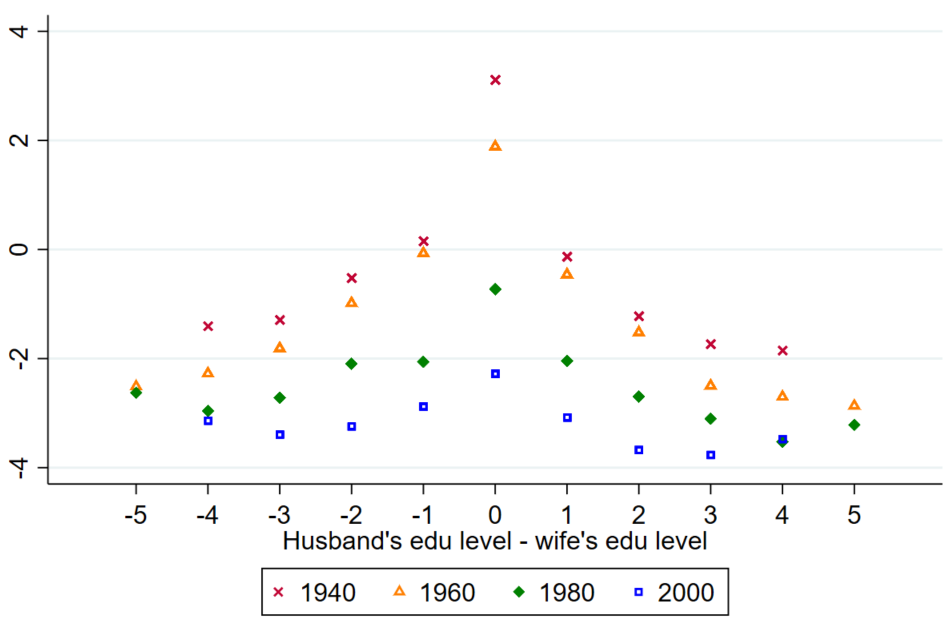


Note: This figure plots the weighted average of labor force participation among men aged between starting age and (starting age+29). It shows that 25-54 is an appropriate age range for the economically active years of one's life and that this observation is quite stable over time.

Figure A4: Histogram of Empirical Market Abilities

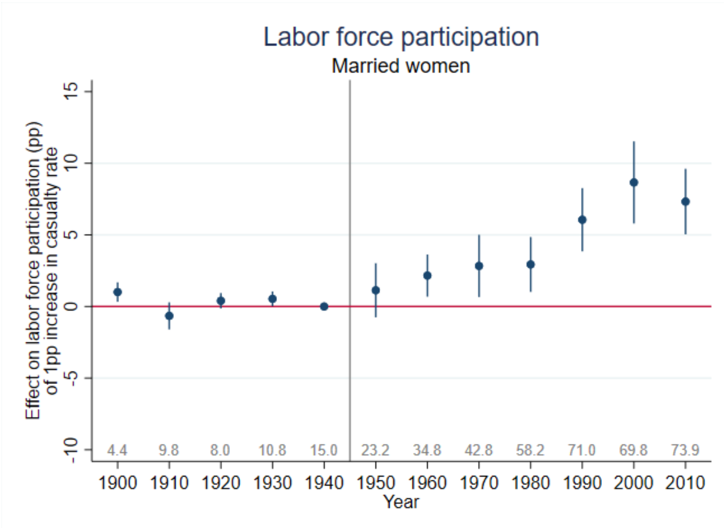
Note: This figure plots the histogram of the market abilities of all working individuals with wage data, where the market abilities are inferred using the model structure as outlined in step 2 of section 4.2.

Figure A5: Ψ by Spousal Education Gap, Over Time

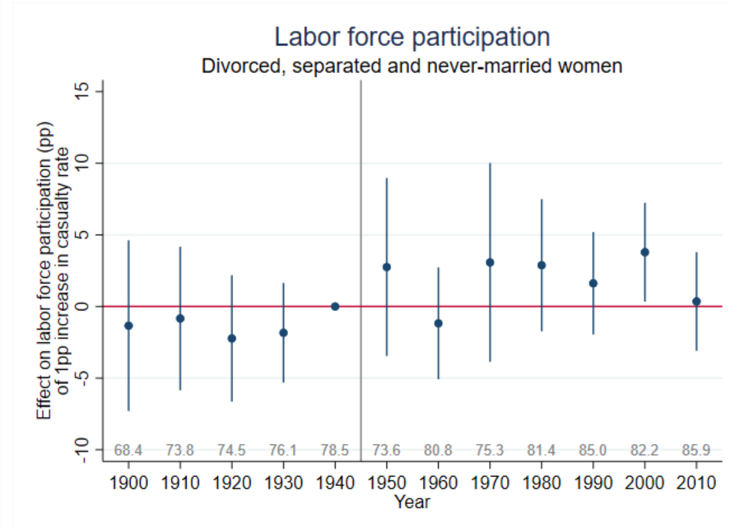


Note: This figure plots the pattern over time of Ψ^{qr}, the utility from marital bliss in a marriage between a man with education level q and a woman with education level r, when averaged by the spousal educational gap $(q - r)$.

Figure A6: The Effect of WW2 Draftee Casualty Rates on Various Outcomes

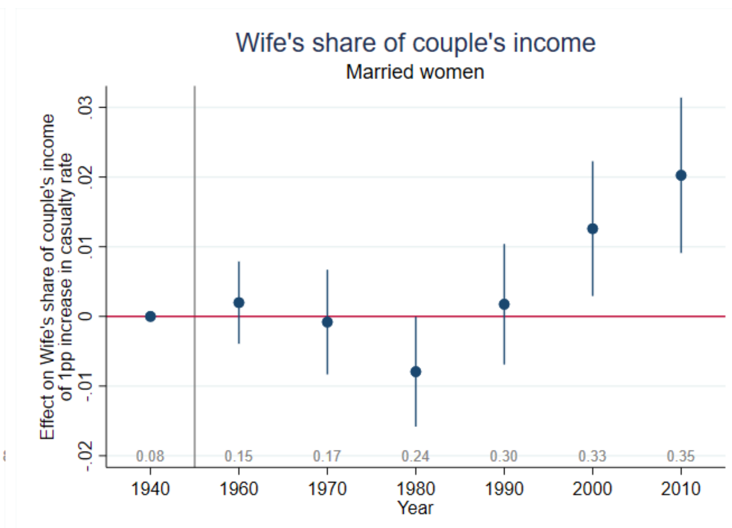


(A)

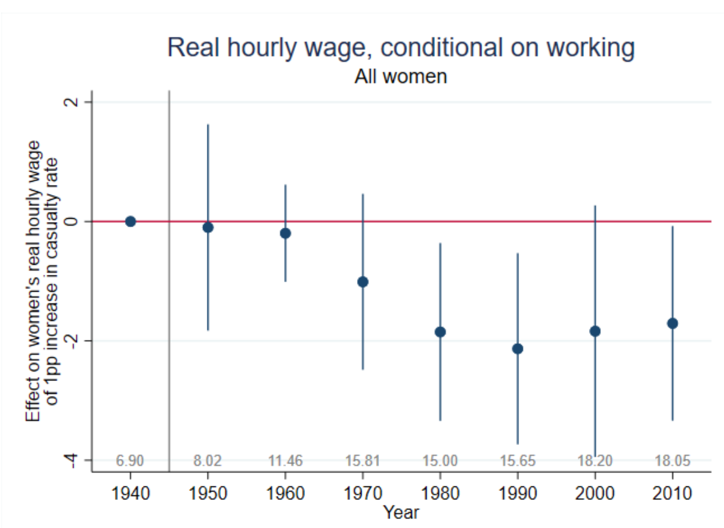


(B)

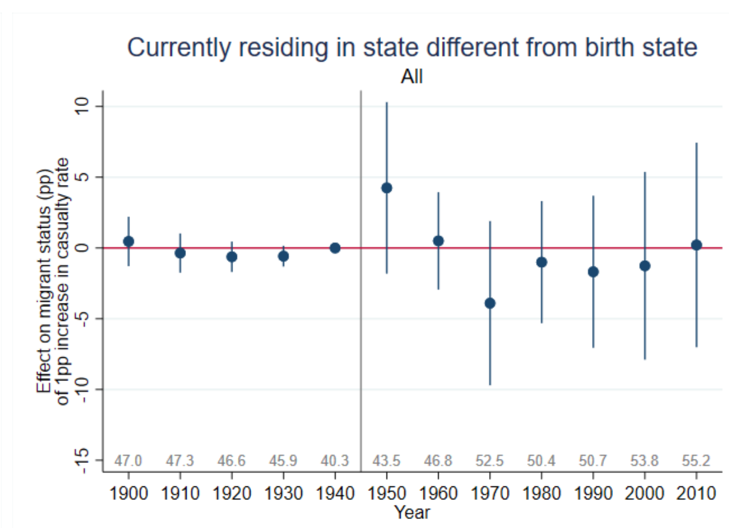
(C)



(D)

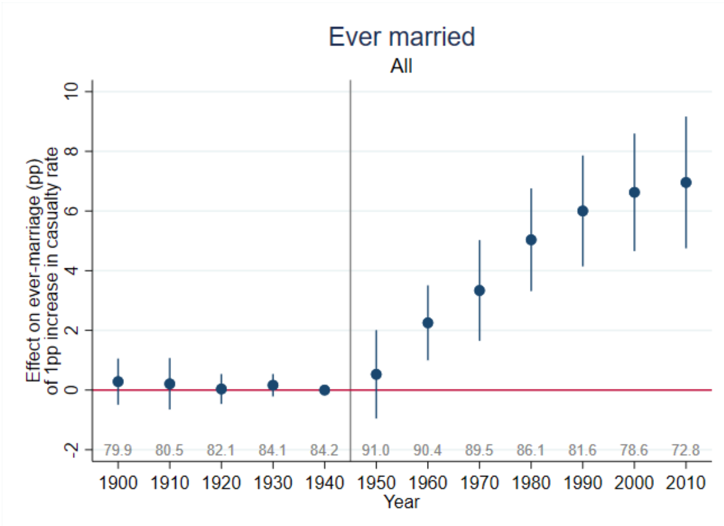


(E)

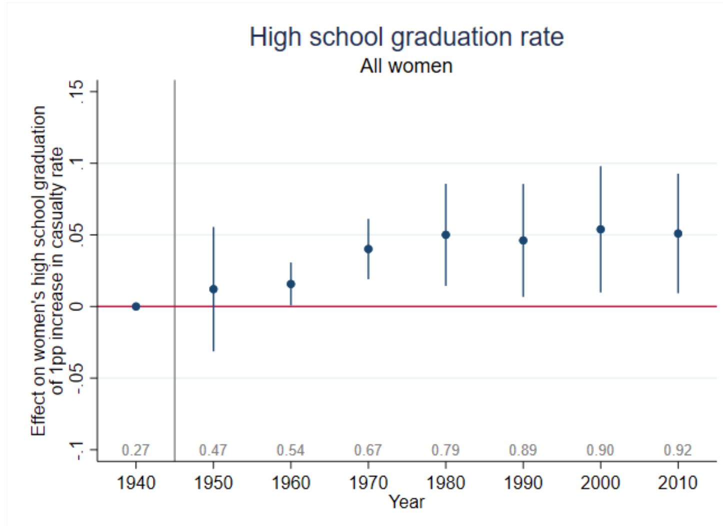


(F)

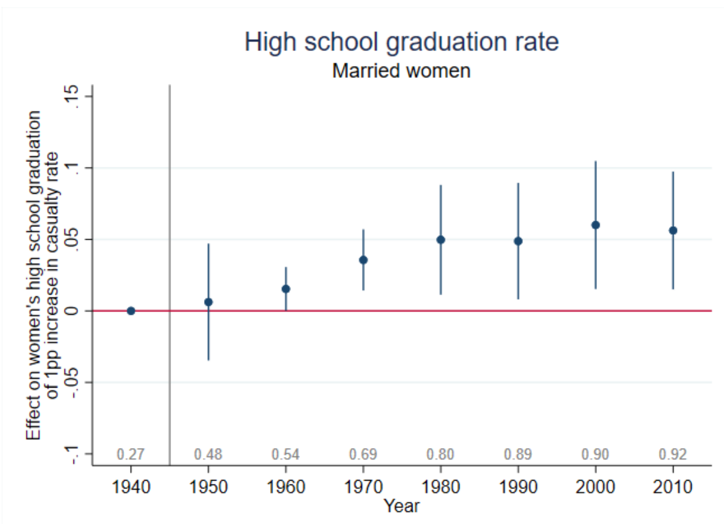
Figure A6 (continued):
The Effect of WW2 Draftee Casualty Rates on Various Outcomes



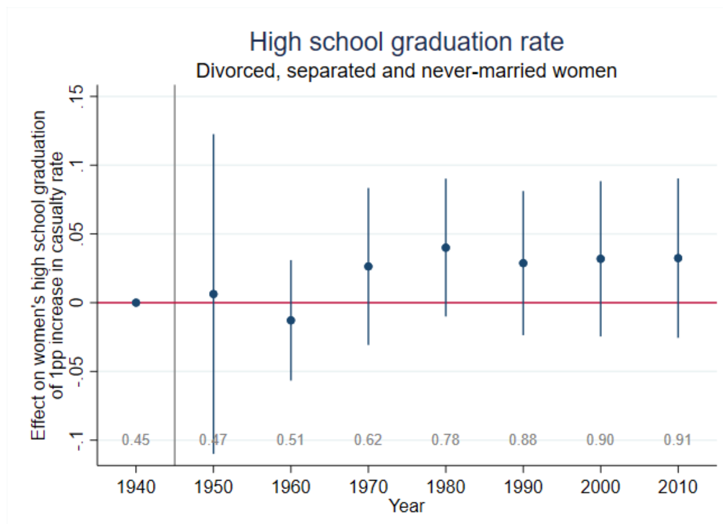
(G)



(H)



(I)



(J)

Note: This figure plots the difference-in-differences coefficients from estimating equation (20) for various outcome variables.

Appendix C Model

C.1 General form of the utility function

The general form of the utility function is given by

$$u_i(Q, C_i, L_f, L_m) = H\Big(f(Q)C_i - r(Q)\tau_i w_f L_f + g_i(Q)\Big)$$

where the following conditions hold:

<u>**Conditions**</u>

C1) H is strictly increasing and strictly concave

C2) $(H')^{-1}$ is homogeneous or logarithmically homogeneous[37,38,39]

C3) $2p(f')^2 - p{\cdot}f{\cdot}f'' + \tau w_f L_f(r''f' - r'f'') - f'g'' + g'f'' > 0$, where $\tau \equiv \tau_m + \tau_f$, and $g(Q) \equiv g_m(Q) + g_f(Q)$

This general utility function yields the same result that 1) the optimal labor supply decision of the couple is one that maximizes their pooled income less the disutilities from the wife working in the labor market, and 2) the optimal labor decisions are made independently based only on individuals' comparisons of the gains from working in the market versus at home.

C.2 Model discussion: On how home production features as "income" in the budget constraint

It might seem non-standard that according to the budget constraint, private and public goods can be "bought" with "income" from home production. However, the maximization

[37]Condition C2) is based on Mazzocco (2004), which shows that a collective household's behavior under uncertainty is equivalent to that of a representative agent if and only if H is of the Identically Shaped Harmonic Absolute Risk Aversion (ISHARA) class:

$$-\frac{H_i''(x)}{H_i'(x)} = \frac{1}{\theta x + a_i}$$

[38]A function is logarithmically homogeneous if it is given by a logarithmic transformation of a homogeneous function. According to Miyake (2015), a function U is logarithmically homogeneous on X if and only if there is a δ-homogeneous function u on X and two parameters $a > 0$ and b such that $U(x) = a \log u(x) + b$ for all $x \in X$. The implication is that $U(\gamma x) = a\delta \log \gamma + U(x)$.

[39]CRRA ($H(x) = \ln(x)$, or $H(x) = \frac{x^{1-\theta}-1}{1-\theta}$ for $\theta > 0, \theta \neq 1$) and CARA ($H(x) = -\exp\{-\theta x + b\}$ for $\theta > 0, b \in R$) utility functions – the most commonly used utility functions for risk-averse individuals – satisfy condition C2).

problem is equivalent to solving

$$\max_{Q,C,Y,B} (Q+Y)\big(C+D-\tau w_f L_f\big)$$
$$\text{s.t.} \qquad pQ + C = w_m L_m + w_f L_f$$
$$pY + B = h_m(1-L_m) + h_f(1-L_f)$$

where Y is the non-rival, public component of home production (e.g. cleaning of communal area, or food preparation for children) and $D \equiv D_m + D_f$ is the total consumption of the private component of home production (e.g. cleaning of private space, laundry of clothes). Here, market goods and services Q and C can only be financed from market earnings, while consumption of home-produced goods and services occur within the total home production done by the couple. The market value of private home goods is normalized to be the same as that of the private market good (the numeraire), and the market value of public home goods is the same as the price of public market good.

C.3 Model discussion: The value of staying home beyond housework

Following Gronau (1977), I can categorize home production into work at home, which is perfectly substitutable to work in the market, and leisure (i.e. home consumption time), which has poor market substitutes. Then, the utility function – in the general form as in Appendix section C.1 – can be formulated as

$$u_i(Q, C_i, L_f, L_m) = H\Big(f(Q)C_i - r(Q)\big[a_i w_i L_i + \tau_i w_f L_f\big] + g_i(Q)\Big)$$

where a_i denotes i's preference for leisure/home consumption time, measured proportionally to i's market wage.[40]

Ultimately, the optimal labor supply decisions would come down to

$$L_m^* = \mathbb{1}\left[w_m - a_m \geq h_m\right]$$
$$L_f^* = \mathbb{1}\left[(1-\tau)w_f - a_f \geq h_f\right]$$

[40] It is easy to also introduce j valuing i's home consumption time, e.g. the husband valuing the wife's play time with their children.

C.4 Model extension: Enriching women's labor supply decisions

The independence in the labor supply decisions of husbands and wives simplifies the study of these decisions. However, a necessary assumption to generate the independence is the perfect substitutability of home-produced goods and market goods. The perfect substitutability assumption, in turn, implies that there are no incentives for specialization. I can enrich the labor supply decisions of women, however, if I abstract from the labor supply decisions of men, i.e. have men always work in the market.

As in Aguiar and Hurst (2007a), consider the final consumption good being produced according to a CES production function with market goods and home production as inputs:

$$\left[\left(y+wL\right)^{\rho}+\left(\underline{h}+h(1-L)\right)^{\rho}\right]^{\frac{1}{\rho}} \tag{22}$$

y refers to nonlabor household market income, which includes any remittances from family, government benefits, and, if married, husband's earnings. $\underline{h}$ refers to the basic level of home production conducted in the household, including the amount of home production a woman would have completed regardless of her labor force participation status, and again, if married, any home production completed by the husband.

To make use of the convenient properties of the Fréchet-distributed market and home abilities, a simplification is necessary. A first-order Taylor approximation of equation (22) around $(w=0, h=0)$ yields

$$(y^{\rho}+\underline{h}^{\rho})^{\frac{1}{\rho}}+(y^{\rho}+\underline{h}^{\rho})^{\frac{1-\rho}{\rho}}y^{\rho-1}wL+(y^{\rho}+\underline{h}^{\rho})^{\frac{1-\rho}{\rho}}\underline{h}^{\rho-1}h(1-L) \tag{23}$$

Hence, the labor supply decision that maximizes equation (23), corresponding to the optimal decision for a single woman, is:

$$\hat{L}^{*}=\mathbb{1}\left[y^{\rho-1}w\geq\underline{h}^{\rho-1}h\right]$$

A married woman, on the other hand, must consider the cost of the norms wedge. Where the cost of working in the market is denominated proportionally to the contribution of market goods in the production of the final good, her optimal labor supply decision is:

$$L^{*}=\mathbb{1}\left[(1-\tau)y^{\rho-1}w\geq\underline{h}^{\rho-1}h\right]$$

The norms wedges from this extension, comparable to ones in the benchmark model given by equation (18), is:

$$\tau^{qr}(\mathcal{K}) = 1 - \frac{\text{avgwage}_F^{0r}(\mathcal{K})}{\text{avgwage}_F^{qr}(\mathcal{K})} \cdot \underbrace{\left(\frac{\underline{h}_F^{0r}}{\underline{h}_F^{qr}} \cdot \frac{y_F^{qr}}{y_F^{0r}} \right)^{1-\rho}}_{\text{new in the extension}} \cdot \left(\frac{1 - P_F^{0r}(\mathcal{K})}{1 - P_F^{qr}(\mathcal{K})} \right)^{\frac{1}{\theta}} \tag{24}$$

The following describes how I compute the additional part of the norms wedge in this extension.

- ρ: set at 0.5, since the elasticity of substitution between market goods and home produced goods in the literature is estimated at around 2; McGrattan, Rogerson, and Wright (1997): 1.8, Chang and Schorfheide (2003): 2.3, Rupert, Rogerson, and Wright (1995): 1.8, Aguiar and Hurst (2007a): 1.78-2.13.

- y: the sum of welfare benefits, social security income, other unearned income, and, if married, spousal earnings in the U.S. Decennial Census. The values are truncated at 0 and winsorized at the top 10% to remove distortions from outliers, and then I take the average by group. I compute (24) only from 1960 because information on non-wage income becomes available from 1960.

- $\underline{h}$: In the benchmark model, single and married women with the same level of education share the same home productivity. Similarly, I assume here that $\frac{\underline{h}_F^{0r}}{\underline{h}_F^{qr}}$ is a constant. As the aim of computing the norms wedge in the extension is to compare its evolution trajectory over time to that of the wedges in the benchmark model, the value of this constant is set to equate the norms wedges in the benchmark and the extension in 1960.

C.5 Model extension: Adding male norms wedges

The independence in the labor supply decisions of husbands and wives, arising from the perfect substitutability of home-produced goods and market goods, allows for the addition of *male* norms wedges. A married couple receives disutility from not only the wife working in the market, but also the husband working in home production. Therefore, married individual $i \in \{m, f\}$ gets the following utility:

$$u_i(Q, C_i, L_f, L_m) = \ln(Q) + \ln\left(C_i - \tau_i w_f L_f - \tilde{\tau}_i h_m(1 - L_m)\right)$$

where $\tilde{\tau}_i$ represents the disutility that i gets from the husband m working. With $\tilde{\tau} \equiv \tilde{\tau}_m + \tilde{\tau}_f$, the optimal labor supply decisions are:

$$L_m^* = \mathbb{1}\left[w_m \geq (1 - \tilde{\tau})h_m\right]$$
$$L_f^* = \mathbb{1}\left[(1 - \tau)w_f \geq h_f\right]$$

C.6 Derivation of marriage market equilibrium

The probability that man m chooses spousal type $r \in \{1, ..., S\}$ or stays single ($r = 0$) is

$$\mathbb{P}(r = \arg\max_{r'=0,1,..,S} V_m^{qr'}) = \frac{\exp\{\mathbb{E}(v_m^{qr}) - \pi^{qr} + \psi^{qr}\}}{\sum_{r'=0}^{S} \exp\{\mathbb{E}(v_m^{qr'}) - \tau^{qr'} + \psi^{qr'}\}}$$

The maximum likelihood estimator of $\mathbb{P}(r = \arg\max_{r'=0,1,..,S} V_m^{qr'})$ is the fraction of type q men married to r, or $\frac{n^{qr}}{M^q}$.

Hence, in terms of the number of (q, r) marriages *demanded* by type q men,

$$\ln n^{qr,\mathcal{D}} = \ln n^{q0} + \mathbb{E}(v_m^{qr}) - \pi^{qr} + \psi^{qr} - \mathbb{E}(\hat{v}_m^{q})$$

Similarly, woman f of type r choosing her spousal type or remaining single gives the analogue for the number of (q, r) marriages *supplied* by type r women,

$$\ln n^{qr,\mathcal{S}} = \ln n^{0r} + \mathbb{E}(v_f^{qr}) + \pi^{qr} + \psi^{qr} - \mathbb{E}(\hat{v}_f^{r})$$

The marriage market-clearing equilibrium transfers π^{qr} are determined such that

$$n^{qr,\mathcal{D}} = n^{qr,\mathcal{S}} = n^{qr}.$$

Appendix D Parameter Inference

D.1 Derivation of the likelihood function to estimate θ

The probability density function of the Fréchet distribution with shape parameter θ and scale parameter s is:

$$f(x; s, \theta) = \frac{\theta}{s} \exp\left\{-\left(\frac{x}{s}\right)^{-\theta}\right\}\left(\frac{x}{s}\right)^{-\theta-1}$$

The scale parameter in the Fréchet distribution of market abilities among market workers is $\left(\frac{1}{P}\right)^{\frac{1}{\theta}}$, where P is the fraction working in the market among a group defined by gender g, match (q, r), and family composition $\mathcal{K}$. Therefore, where x_n is the market ability of observation n and P_n denotes the fraction of workers in observation n's group, the maximum likelihood estimator for θ is:

$$\tilde{\theta}_{MLE} = \arg \max_{\theta \in (0,\infty)} \sum_{n=1}^{Nobs} \left[\ln \theta - \ln P_n - x_n^{-\theta} P_n^{-1} - (\theta + 1) \ln x_n\right]$$

D.2 Extracting market abilities from market wages

Recall that an individual i of gender g in a (q, r) match with family composition $\mathcal{K}$ receives wage

$$w_{it} = \bar{w}_{gt}^{qr}(\mathcal{K})\epsilon_i^w, \quad g \in \{M, F\}$$

where $\bar{w}_{gt}^{qr}(\mathcal{K})$ is the market wage per unit of effective labor for each group. Therefore the x_n in practice equals ϵ_{it}^{w*} where the asterisk denotes that this it the market ability of those who choose to work in the market. Let us isolate ϵ_i^{w*} from the observed wages:

$$\text{logwage}_{it} = \ln \bar{w}_{gt}^{qr}(\mathcal{K}) + \ln \epsilon_{it}^{w*}$$

To this end, I regress log wages on (decade × sex × education pair × family composition) dummies. For each group, then, the residuals are

$$\begin{aligned} \text{residuals}_{it} &= \text{logwage}_{it} - \overline{\text{logwage}}_{it} \\ &= \left(\ln \bar{w}_{gt}^{qr}(\mathcal{K}) + \ln \epsilon_{it}^{w*}\right) - \left(\ln \bar{w}_{gt}^{qr}(\mathcal{K}) + \overline{\ln \epsilon_{it}^{w*}}\right) \\ &= \ln \epsilon_{it}^{w*} - \overline{\ln \epsilon_{it}^{w*}} \end{aligned}$$

Thus,

$$\begin{aligned} x_n &= \exp\{\text{residuals}_{it} + E(\ln \epsilon_{it}^{w*})\} \\ &= \exp\{\text{residuals}_{it} + \tfrac{\gamma}{\theta} + \ln(s_n)\} \\ &= \exp\{\text{residuals}_{it} + \tfrac{\gamma}{\theta} - \tfrac{1}{\theta}\ln P_n\} \end{aligned}$$

where γ is the Euler's constant.[41]

D.3 Reformulation of the expected utility from each schooling level

From the marriage market clearing condition and equation (4),

$$\begin{aligned} \pi^{qr} &= \frac{\ln n^{q0} - \ln n^{0r} - \ln \mu^{qr} - \mathbb{E}(\hat{v}_m^q) + \mathbb{E}(\hat{v}_f^r)}{2} \\ &= \frac{\ln n^{q0} - \ln n^{0r} - \ln \mu^{qr} - 2\mathbb{E}(\ln \hat{Q}_m^q) + 2\mathbb{E}(\ln \hat{Q}_f^r)}{2} \end{aligned}$$

Then,

$$\begin{aligned} U_F^r &= \sum_{q=0}^{S}\left[\frac{n^{qr}}{F^r}\Big(\mathbb{E}(v_f^{qr}) + \pi^{qr} + \psi^{qr}\Big)\right] - c_F^r - \xi^r \\ &= \sum_{q=0}^{S}\left[\frac{n^{qr}}{F^r}\left(2\mathbb{E}(\ln Q^{qr}) + \ln\frac{\mu^{qr}}{1+\mu^{qr}} + \pi^{qr} + \psi^{qr}\right)\right] + \ln p - c_F^r - \xi^r \\ &= \sum_{q=0}^{S}\left[\frac{n^{qr}}{F^r}\left(2\mathbb{E}(\ln Q^{qr}) + \frac{1}{2}\ln \mu^{qr} - \ln(1+\mu^{qr}) + \frac{\ln n^{q0} - \ln n^{0r}}{2} - \mathbb{E}(\ln \hat{Q}_m^q) + \mathbb{E}(\ln \hat{Q}_f^r) + \psi^{qr}\right)\right] \\ &\qquad + \ln p - c_F^r - \xi^r \\ &= \sum_{q=0}^{S}\left[\frac{n^{qr}}{F^r}\left(2\mathbb{E}(\ln Q^{qr}) + \Psi^{qr} + \frac{\ln n^{q0} - \ln n^{0r}}{2} - \mathbb{E}(\ln \hat{Q}_m^q) + \mathbb{E}(\ln \hat{Q}_f^r)\right)\right] + \ln p - c_F^r - \xi^r \end{aligned}$$

[41] To compute $E(\ln \epsilon_{it}^{w*})$, I need the probability density function of $y = g(x) = \ln(x)$ where x is a Fréchet random variable with scale parameter s and shape parameter θ. $f_Y(y) = f_X(g^{-1}(y))\left|\frac{dx}{dy}\right| = \frac{\theta}{s}\left(\frac{e^y}{s}\right)^{-\theta-1} e^{-\left(\frac{e^y}{s}\right)^{-\theta}}|e^y|$. Thus, $E(y) = \theta s^\theta \int_{-\infty}^{\infty} y e^{-\theta y} e^{-e^{-\theta y}s^\theta} dy = -\frac{1}{\theta}\int_0^\infty e^{-z}(\ln z - \theta \ln s)dz = -\frac{1}{\theta}\Gamma'(1) + \ln s = \frac{\gamma}{\theta} + \ln s$ where γ is the Euler's constant.

Appendix E Norms Wedges under Alternative Models

E.1 Male norms wedges

What if married *men* are also subject to norms wedges? The male counterpart to the homemaker gender role would be the breadwinner gender role where married men are expected to earn a market income. This situation corresponds to the model extension in Appendix section C.5. Because the optimal labor supply decisions of husbands and wives are independent, the values of the norms wedges applying to married women are not affected. Nonetheless, it would be interesting to compare the "female" norms wedges to the "male" norms wedges.

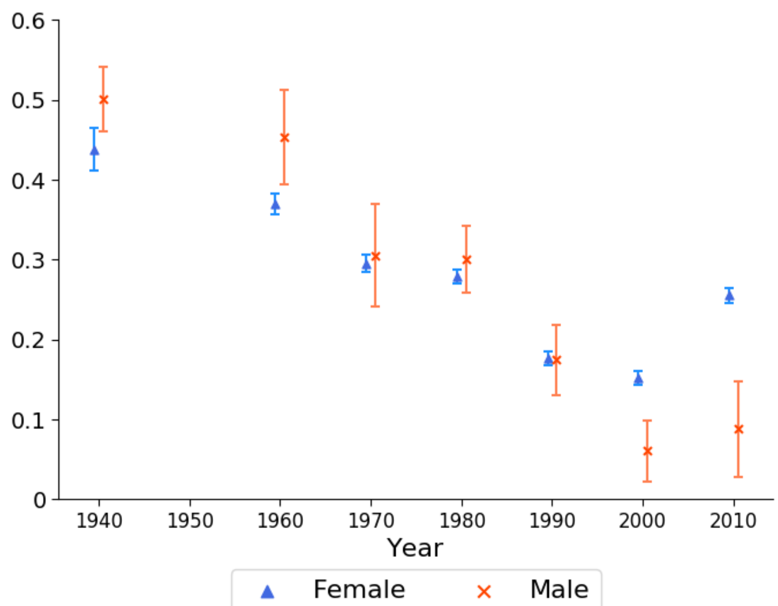


Figure A7: Female and Male Norms Wedges by Decade

Note: This figure plots the weighted median of τ and $\tilde{\tau}$ (described in Appendix section C.5), inferred for each group, by decade. The weight equals the empirical probability of each group. The error bars indicate 95% confidence intervals based on bootstrapped standard errors with 50 replications.

Figure A7 shows that the male norms wedge declines over time, resembling the trend of the female norms wedge. The decline indicates that the labor force participation of married men are falling faster than that of comparable single men. In fact, the male norms wedge is even higher than the female counterpart in 1940 and is lower towards the end of the sample period. This result may be counterintuitive; it is difficult to reconcile the consistent and large decline with the fact that the rise of the stay-at-home fathers is only a recent phenomenon.

However, the male norms wedges must be interpreted with a grain of salt. The first reason is an algebraic one. The male norms wedges are proportional to the ratio of *nonparticipation in the labor force* of married and single men.[42] Throughout the sample period, the male labor force participation rate is very close to 1 regardless of marital status, so that small changes in the labor force participation rate translate to large changes in norms wedges. For the same reason, the male norms wedge is also much noisier than the female counterpart, as depicted by the wider confidence intervals.[43] The second reason is that nonparticipation in the labor force is less synonymous with home production for men, so $\tilde{\tau}$ would have less to do with actual gender norms. Indeed, state-level male norms wedges are *not* correlated to state-level attitudes towards gender roles, unlike female norms wedges (Table 2).

E.2 Robustness of norms wedges

This section demonstrates the robustness of the values of the norms wedges to alternative model specifications and sample selection criteria.

Figure A8 plots the norms wedges under four different specifications. "Baseline" refers to the norms wedges in the baseline model, and is a reproduction of the norms wedges in Figure 7. "With income taxes" is a modification of the baseline model where individuals compare home productivity to not *before*-tax market wages but *after*-tax market wages. Another way in which a married woman's labor supply decision might differ from a similar single woman's is the aspect of social insurance, where marriage may allow individuals to save on taxes paid. Thus, it is important to check for robustness to after-tax earnings. Because precise after-tax earnings are not recorded in the U.S. Decennial Census, I apply to the reported before-tax earnings, the U.S. federal individual nominal income tax rates from the Tax Foundation (2013). I assume that individuals file for taxes under the most profitable category that they are eligible for, among Married Filing Jointly, Married Filing Separately, Single, and Head of Household.

"Baseline, ages 40-54 only" refers to the norms wedges in the baseline model where the sample excludes 25- to 39-year-olds. This sample selection criteria addresses the concern that a married woman's labor force participation may differ from a similar single woman's because she has different child-bearing prospects. Child-bearing prospects matter for a woman's

[42]The male norms wedge for a married man in a (q, r) match with family composition $\mathcal{K}$ is

$$\tilde{\tau}^{qr}(\mathcal{K}) = 1 - \frac{\text{avgwage}^{qr}_M(\mathcal{K})}{\text{avgwage}^{q0}_M(\mathcal{K})} \left(\frac{1 - P^{qr}_M(\mathcal{K})}{1 - P^{q0}_M(\mathcal{K})} \right)^{\frac{1}{\theta}}$$

[43]The noisiness can be viewed as a consequence of the difficulty of estimation when the parameter is near the boundary (Andrews, 1999).

decision to work in the labor market. For example, with positive returns to experience, a woman would be more likely to work if she expects fewer interruptions during her career cycle. Child-bearing prospects also matter from the employer's perspective; an employer may be less willing to hire a woman who they view as likely to leave temporarily or permanently in the future. Hence, I restrict the sample to women who are less likely to conceive more children.

It is reassuring that for all three specifications, the values and trends of the norms wedges are similar.

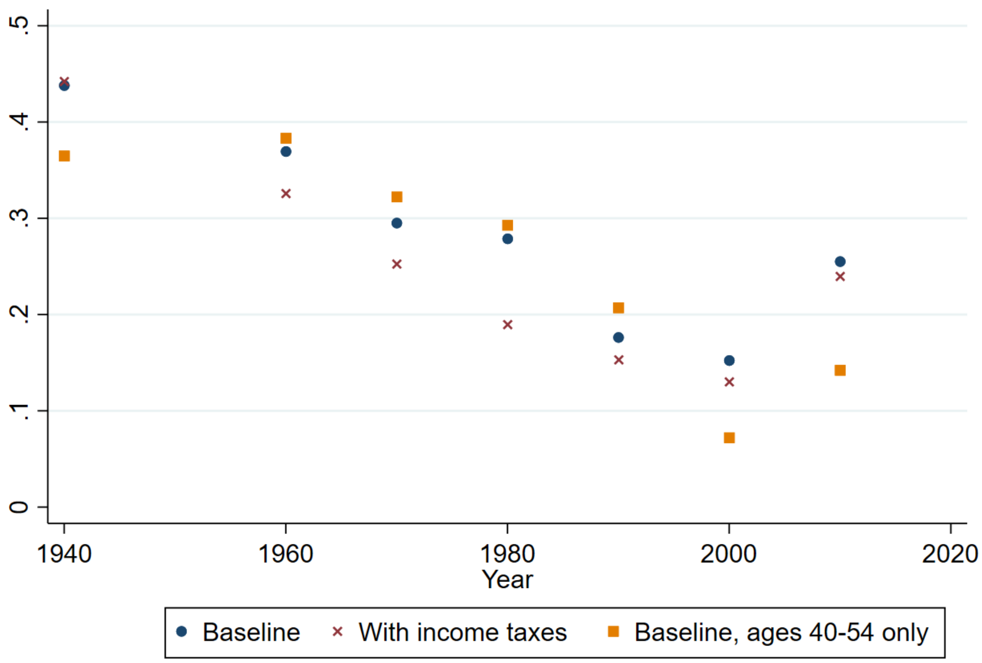


Figure A8: Norms Wedges Under Alternative Specifications or Sample Criteria

Note: This figure plots the values of the norms wedges by decade, under alternative model specifications or sample criteria. "Baseline" refers to the norms wedges in the baseline model, and is a reproduction of the norms wedges in Figure 7. "With income taxes" is a modification of the baseline model where individuals compare home productivity to not *before*-tax market wages but *after*-tax market wages. The tax rates are based on the U.S. federal individual nominal income tax rates from the Tax Foundation (2013). "Baseline, ages 40-54 only" refers to the norms wedges in the baseline model where the sample excludes 25- to 39-year-olds.

References

Acemoglu, D., D. H. Autor, and D. 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.

Aguiar, M. and E. Hurst (2007a). Life-cycle prices and production. *American Economic Review 97*(5), 1533–1559.

Aguiar, M. and E. Hurst (2007b). Measuring Trends in Leisure: The Allocation of Time Over Five Decades*. *The Quarterly Journal of Economics 122*(3), 969–1006.

Akerlof, G. A. and R. E. Kranton (2000). Economics and Identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Albanesi, S. and C. Olivetti (2016). Gender roles and medical progress. *Journal of Political Economy 124*(3), 650 – 695.

Andrews, D. W. K. (1999). Estimation when a parameter is on a boundary. *Econometrica 67*(6), 1341–1383.

Arkhangelsky, D., S. Athey, D. A. Hirshberg, G. W. Imbens, and S. Wager (2021). Synthetic Difference In Differences. *American Economic Review* (forthcoming).

Bar, M., O. Leukhina, M. Hazan, and H. Zoabi (2018). Why did Rich Families Increase their Fertility? Inequality and Marketization of Child Care. *Journal of Economic Growth 23*, 427–463.

Becker, G. (1981). *A Treatise on the Family.* Cambridge: Harvard University Press. Enlarged Edition, 1991.

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

Bedi, A. S., T. Majilla, and M. Rieger (2021). Does signaling childcare support on job applications reduce the motherhood penalty? *Review of Economics of the Household*.

Bergstrom, T. C. and R. C. Cornes (1983). Independence of allocative efficiency from distribution in the theory of public goods. *Econometrica 51*(6), 1753–1765.

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender Identity and Relative Income within Households. *The Quarterly Journal of Economics 130*(2), 571–614.

Bisin, A. and T. Verdier (2000). "Beyond the Melting Pot": Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits*. *The Quarterly Journal of Economics 115*(3), 955–988.

Blau, F. D. and L. M. Kahn (2007). Changes in the labor supply behavior of married women: 1980–2000. *Journal of Labor Economics 25*(3), 393–438.

Blundell, R., M. Graber, and M. Mogstad (2015). Labor income dynamics and the insurance from taxes, transfers, and the family. *Journal of Public Economics 127*(C), 58–73.

Browning, M., P.-A. Chiappori, and Y. Weiss (2014). *Economics of the Family*. Cambridge University Press.

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2020, October). Misperceived social norms: Women working outside the home in saudi arabia. *American Economic Review 110*(10), 2997–3029.

Case, A. and A. Deaton (2017). Mortality and Morbidity in the 21st Century. *Brookings Papers on Economic Activity, Spring 2017*.

Chetty, R. (2006). A new method of estimating risk aversion. *The American Economic Review 96*(5), 1821–1834.

Chiappori, P.-A., M. Costa-Dias, and C. Meghir (2018). The Marriage Market, Labor Supply, and Education Choice. *Journal of Political Economy 126*(S1), 26–72.

Chiappori, P.-A., M. Iyigun, and Y. Weiss (2009). Investment in Schooling and the Marriage Market. *American Economic Review 99*(5), 1689–1713.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2017). Partner Choice, Investment in Children, and the Marital College Premium. *American Economic Review 107*(8), 2109–67.

Choo, E. and A. Siow (2006). Who Marries Whom and Why. *Journal of Political Economy 114*(1), 175–201.

Couprie, H., E. Cudeville, and C. Sofer (2020). Efficiency versus gender roles and stereotypes: an experiment in domestic production. *Experimental Economics 23*(1), 181–211.

Doepke, M., M. Hazan, and Y. D. Maoz (2015). The Baby Boom and World War II: A Macroeconomic Analysis. *Review of Economic Studies 82*(3), 1031–1073.

Doepke, M. and M. Tertilt (2016). Families in macroeconomics. Volume 2, Chapter Chapter 23, pp. 1789–1891. Elsevier.

Eaton, J. and S. Kortum (2004). Technology, Geography, and Trade. *Econometrica 70*(5), 1741–1779.

Erosa, A., L. Fuster, G. Kambourov, and R. Rogerson (2022, July). Hours, Occupations, and Gender Differences in Labor Market Outcomes. *American Economic Journal: Macroeconomics 14*(3), 543–590.

Fernández, R. (2013). Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century. *American Economic Review 103*(1), 472–500.

Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and Sons: Preference Formation and Female Labor Force Dynamics*. *The Quarterly Journal of Economics 119*(4), 1249–1299.

Fernández, R. and J. Wong (2014). Unilateral Divorce, the Decreasing Gender Gap, and Married Women's Labor Force Participation. *American Economic Review 104*(5), 342–47.

Ferrara, A. (2021). World War II and Black Economic Progress. *Journal of Labor Economics* (just accepted).

Field, E., R. Pande, N. Rigol, S. Schaner, and C. Troyer Moore (2021, July). On her own account: How strengthening women's financial control impacts labor supply and gender norms. *American Economic Review 111*(7), 2342–75.

Fogli, A. and L. Veldkamp (2011). Nature or Nurture? Learning and the Geography of Female Labor Force Participation. *Econometrica 79*(4), 1103–1138.

Gayle, G.-L. and A. Shephard (2019). Optimal taxation, marriage, home production, and family labor supply. *Econometrica 87*(1), 291–326.

Goldin, C. (1995). The U-Shaped Female Labor Force Function in Economic Development and Economic History. In T. P. Schultz (Ed.), *Investment in Women's Human Capital and Economic Development*, pp. 61–90. University of Chicago Press.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review 96*(2), 1–21.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.

Goldin, C. and L. F. Katz (2002). The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions. *Journal of Political Economy 110*(4), 730–770.

Goldin, C. D. (1991). The Role of World War II in the Rise of Women's Employment. *American Economic Review 81*(4), 741–756.

Goussé, M., N. Jacquemet, and J.-M. Robin (2017). Marriage, labor supply, and home production. *Econometrica 85*(6), 1873–1919.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2016). Technology and the Changing Family: A Unified Model of Marriage, Divorce, Educational Attainment, and Married Female Labor-Force Participation, journal = American Economic Journal: Macroeconomics. *8*(1), 1–41.

Greenwood, J., A. Seshadri, and M. Yorukoglu (2005). Engines of Liberation. *The Review of Economic Studies 72*(1), 109–133.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica 45*(1), 1–22.

Gronau, R. (1977). Leisure, home production, and work-the theory of the allocation of time revisited. *Journal of Political Economy 85*(6), 1099–1123.

Guner, N., R. Kaygusuz, and G. Ventura (2012). Taxation and household labour supply. *Review of Economic Studies 79*(3), 1113–1149.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2019). The Allocation of Talent and U.S. Economic Growth. *Econometrica 87*(5), 1439–1474.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and Manufacturing TFP in China and India*. *The Quarterly Journal of Economics 124*(4), 1403–1448.

International Labor Organization (2019). Labor force participation rate, female (% of female population ages 15+). International Labor Organization Database (ILOSTAT).

Juhn, C. and K. McCue (2017). Specialization Then and Now: Marriage, Children, and the Gender Earnings Gap across Cohorts. *Journal of Economic Perspectives 31*(1), 183–204.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics 11*(4), 181–209.

Kleven, H. J., C. T. Kreiner, and E. Saez (2009). The optimal income taxation of couples. *Econometrica 77*(2), 537–560.

Kuziemko, I., J. Pan, J. Shen, and E. Washington (2018). The mommy effect: Do women anticipate the employment effects of motherhood? Working Paper 24740, National Bureau of Economic Research.

Lee, J. E. (2021). Homemaker Norms for Married Women: A Cross-Country Comparison. Working paper.

Lee, U. (1965). The Employment of Negro Troops. In S. Conn (Ed.), *United States Army in World War II*. Washington D.C: Center of Military History U.S. Army.

Lundberg, S. and R. A. Pollak (2007). The american family and family economics. *Journal of Economic Perspectives 21*(2), 3–26.

Mammen, K. and C. Paxson (2000). Women's work and economic development. *Journal of Economic Perspectives 14*(4), 141–164.

Mazzocco, M. (2004). Saving, Risk Sharing, and Preferences for Risk. *American Economic Review 94*(4), 1169–1182.

Michalopoulos, S. and M. M. Xue (2021, 01). Folklore*. *The Quarterly Journal of Economics 136*(4), 1993–2046.

Miyake, M. (2016). Logarithmically homogeneous preferences. *Journal of Mathematical Economics 67*, 1 – 9.

Ngai, L. R. and B. Petrongolo (2017). Gender Gaps and the Rise of the Service Economy. *American Economic Journal: Macroeconomics 9*(4), 1–44.

Ngai, R., C. Olivetti, and B. Petrongolo (2020). Structural Transformation and the U-Shaped Female Labor Supply. Working paper.

Pollak, R. A. (2013). Allocating Household Time: When Does Efficiency Imply Specialization? NBER Working Papers 19178, National Bureau of Economic Research, Inc.

Restuccia, D. and R. Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics 11*(4), 707–720.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers 3*(2), 135–146.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2021). IPUMS USA: Version 11.0 [dataset]. `https://doi.org/10.18128/D010.V11.0`. Minneapolis, MN: IPUMS.

Shapley, L. and M. Shubik (1971). The Assignment Game I: The Core. *International Journal of Game Theory 1*(1), 111–130.

Tax Foundation (2013). U.S. Federal Individual Income Tax Rates History, 1862-2013 - Nominal Dollars [table].