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

Jay Euijung Lee†

JOB MARKET PAPER

November 14, 2019 – Latest version here

Abstract

By how much do traditional gender norms in marriage constrain aggregate output? Married women are traditionally expected to stay home and take care of the household. This gender role reduces married women's labor force participation, away from their comparative advantage. A low likelihood of working in the future also reduces women's incentive to get educated. I develop a model featuring education, marriage, and labor supply choices to quantify the aggregate economic consequences of gender norms in marriage. I find that relative to single women, married women in 1940 U.S. behaved as if they faced a 44% "norms tax" on market wage. By 2010, the norms tax had halved. Had gender norms remained at the level of 1940, married women of 2010 would have had an 18% lower labor force participation rate, 13% lower market earnings, and their total market and home output would have been lower by 7%. For the aggregate economy, total market and home output would have been 3.5% lower. I validate the model structure through a reduced form analysis, which uses county-level variation in World War 2 casualties that increased female labor force participation and consequently weakened traditional gender norms.

JEL codes: I25, J12, J16, J24, O47, O51

Keywords: gender roles, marriage, misallocation, labor force participation, aggregate output

*Special thanks to my advisors, Maitreesh Ghatak, Guy Michaels, Oriana Bandiera, and Robin Burgess, for their invaluable guidance, support, and patience. I am very grateful to Gharad Bryan and Tim Besley for countless constructive feedback and generosity with their time. For very helpful comments, I thank Andrea Alati, Doug Allen, Alexandra Benham, Lee Benham, Francesco Caselli, Alexia Delfino, Greg Fischer, Alessandro Gavazza, Ethan Ilzetski, Chad Jones, Emir Kamenica, Camille Landais, Gary Libecap, Stephen Machin, Alan Manning, Ameet Morjaria, Rachel Ngai, John Nye, Barbara Petrongolo, Steve Pischke, Gautam Rao, John Van Reenen, Maros Servatka, David Skarbek, John Sullivan, Silvana Tenreyro, Scott Wallsten, Horng Chern Wong, seminar participants at LSE and participants in RCI Institutional Analysis Workshop and RHUL PhD conference 2019. 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**.

†Department of Economics, London School of Economics; email: E.Lee13@lse.ac.uk.

1 Introduction

Many women shift their time from the labor market to home production upon marriage (Lundberg and Pollak, 2007).[1] 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 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 market.

By how much do traditional gender norms in marriage constrain aggregate output? This paper aims to quantify 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., 2019; Bedi et al., 2018; Couprie et al., 2017), little is known about the aggregate implications of gender norms. Existing papers on the aggregate implications of gender differences such as labor market and educational market discrimination (Hsieh et al., 2019) and nonmarket time (Erosa et al., 2017) do not distinguish between *married and single* women, despite markedly disparate labor market outcomes (Goldin, 2006; Blau and Kahn, 2007). I contribute by focusing on gender roles associated with *marriage*.

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 single women, besides wage differentials. The model structure is validated by a reduced form analysis, which uses county-level variation in World War 2 casualties that increased female labor force participation and consequently weakened traditional gender norms. 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 change over time. With the model, I then compute counterfactuals to determine by how much aggregate output is affected by changing gender roles and dissect the underlying channels.

My central finding is that gender roles have changed significantly in the U.S. and that gender norms have important output effects. If gender norms had stayed at the level of 1940, aggregate market output in 2010 would be lower by 4.8% and aggregate total (market and nonmarket) output would be lower by 3.5%. Gender norms matter more for the subpopulation of married women, whose labor force participation is directly affected. Married women in this counterfactual cumulatively have a 13.0% lower market output and 6.5% lower total output.

To establish these findings, I start by presenting empirical facts that motivate my focus on traditional gender roles as the distinguishing factor between married and single women. I use the U.S. decennial census to establish the first motivating fact: married women's labor force participation that is *not* accounted for by standard observables, such as age, education, race, and the number of

[1] I show this explicitly later in section 2.

children, rises over time to catch up with the stable counterpart for single women. This disparate trend highlights the importance of unobservable variables as drivers of married women's labor force participation, including traditional gender roles that affect married women but not single women. The second motivating fact, derived from the Panel Survey of Income Dynamics (1968-2015), is that individuals undergo stark changes in their time use right in the first year of marriage. The share of housework hours[2] relative to paid work hours falls sharply for men and rises sharply for women getting married in the 1970s. For later marriages, however, there are no such sharp changes in time use upon marriage. This finding highlights that being married shifts the responsibility of house chores to women, but in a way that weakens over time.

With these motivating facts in mind, I develop a structural model for two purposes. I first use it as a measurement tool to quantify by how much gender roles affect married women's labor supply choice, featured as a parameter in the model. Then I use the model to conduct counterfactuals to gauge the importance of gender roles in marriage for various aggregate measures related to labor supply, earnings, marriage, education, and most importantly, output.

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. Second, they enter the marriage market, a frictionless transferable utility set up 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 draws a family composition category (e.g. 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 in home production for each individual. Gender roles are modeled as a disutility that a married couple gets when the wife works in the market. 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 set up, gender roles directly affect labor supply choice, and also indirectly affect marriage and educational choice in anticipation.

My theoretical contribution is fourfold. Firstly, I augment a tractable form of selection into labor activity by individuals with heterogeneous abilities, derived from the Trade literature (Eaton and Kortum, 2004; Hsieh et al., 2019), with concerns over fulfilling gender roles (Akerlof and Kranton, 2000; Bertrand, Kamenica, and Pan, 2015). Secondly, I embed this form of selection, previously used to study *individual* choices, into a model of *household* decision-making. Thirdly, I ensure that the household economic utilities resulting from the labor activity choice are fully consistent with models of educational choice and marriage market matching (Choo and Siow, 2006; Chiappori, Salanié, and Weiss, 2017; Chiappori, Dias, and Meghir, 2018). Lastly, the recursive structure of my model simplifies the parameter identification procedure and allows me to manage a very large number of household types.

I calibrate the model to match the education, marriage, labor force participation patterns in

[2] The answer to the question, "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?"

the U.S. decennial census decade by decade, assuming that the data is a reflection of the model equilibrium. The practical advantage of my model is that it is not demanding on the data, as the only variables needed are market wage, labor force participation, marital status, education and children, all of which become available from 1940.[3] As the model is fitted decade by decade, model parameters other than norms wedges flexibly account for secular changes in the gender wage gap, gender differences in home productivity, propensity of marriage, assortativeness of marriage matching by education, and educational attainment.

I find that married women faced a 44% norms wedge on the market wage in 1940, which declined to 25% by 2010. To cross-check whether these wedges correlate with more conventional measures of gender norms, I repeat the calibration at the *state level* and regress state-level averages of the norms wedges on the state-level average answers to attitudinal surveys related to gender roles in marriage.[4] I find that the states that answer more conservatively in attitudinal surveys are also the ones with higher norms wedges.

I use the model to conduct a counterfactual, where I consider what would have happened in 2010, had gender norms not changed since 1940. I first find that the number of completed school years of women drops by 1.4% and that of men by 0.8%. This is the result of marriage becoming less attractive and thus the marriage-market returns to education falling. The effect on women is compounded by falling labor-market returns to education. The marriage rate indeed falls by 32.2%. As higher-earning women are affected by more with higher norms wedges, the selection into marriage by education becomes more negative for women. Due to assortative matching on education, the selection into marriage by education becomes more negative for men as well. Moreover, married women's labor force participation and cumulative market earnings drop by 17.5% and 13.0%, respectively. On aggregate, the total market output of the economy falls by 4.8% and the total market *and* nonmarket output drops by 3.5%. The effect on total output amounts to a half of the drop in total output that would be seen if married women of 2010 made labor force participation choices based on 1940 wages and home productivities. The finding of smaller effects on output than on labor force participation echoes Hsieh et al. (2019)'s findings on the aggregate effects of occupation-specific 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 empirical trends in the United States. Specifically, these are a) rising married female labor force participation rate, b) rise in wife's share of household income, c) faster growth of educational attainment of women relative to men, and d) increasingly positive selection into marriage by education of both men and women (Bar et al., 2018; Juhn and McCue, 2016; Case and Deaton, 2017).

Since the counterfactual results depend on the model structure, I next perform a reduced form

[3]This feature of the model implies that the model can be easily applied to many other settings.

[4]I put together various surveys in the Roper Polls Database. Questions include whether one approves of a married woman working if she has a husband capable of supporting her and whether it is more important for a wife to help her husband's career than to have one herself, and many others.

exercise to validate the model. 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 in a way that indicates 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 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 structural model to have norms evolving dynamically in response to past female labor force participation. I augment the model to have *economywide* norms evolve in response to *economywide* female labor force participation in the past decade, a relationship that I estimate based on the reduced form coefficients. This extension of the model is compatible with how I identified norms wedges previously, as long as individuals take norms as given and do not internalize the effect of their labor supply choice on the norms of future generations. The model extension further enables me to conduct *dynamic* counterfactuals, on how a shock would affect an economy *over time*. A simple thought experiment of females temporarily getting paid male wages in 2010, shows that the economy of 2010 stabilizes within three decades at 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,[5] 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). Thus, this paper can be informative 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.

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 misal-

[5]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.

location. 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. (2017), 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 that are associated with marriage, and *quantifying* 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. (2017) 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 female 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 remainder of the paper is organized as follows. Section 2 describes empirical facts that motivate my focus on the distinction between married and single women's labor supply decisions. Section 3 sets up the structural model, making explicit this difference. Section 4 describes the data and how model parameters are calibrated to fit the model to the data. It then discusses the calibration results. Section 5 quantifies the effect of changes in gender norms through the lens of the model and benchmarks the results to the effects of other comparable counterfactuals. 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 investigation of the aggregate output effects of gender roles in marriage.

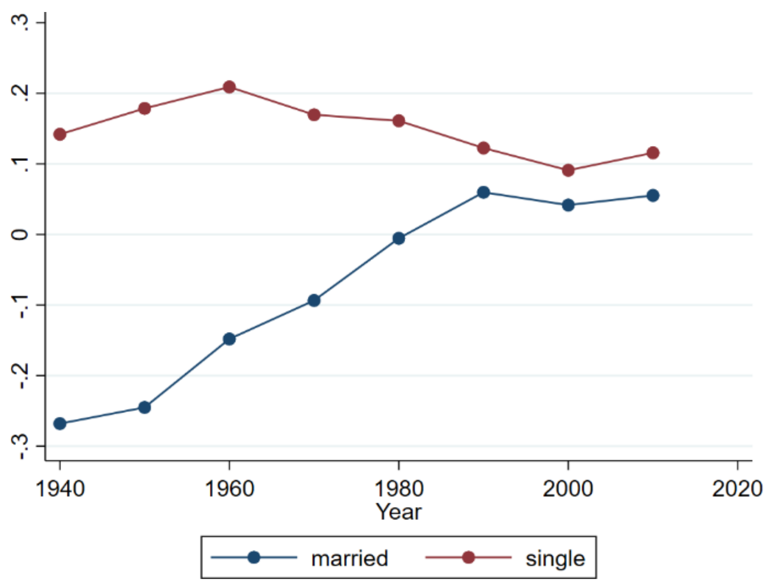
2.1 Married vs. Single "Unexplained" Female Labor Force Participation

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 labor force participation status indicator Lab on standard, commonly observed individual characteristics X: age, education, race, and the number of children dummies.

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

The regression sample is all females aged 25-54 between 1940 and 2010 in the U.S. decennial census. Then I take the weighted average of the residuals by marital status and decade.

Figure 1: Residualized female labor force participation, by marital status and decade



Notes: 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 the number of children dummies. See text for the residualization procedure.

Figure 1 shows that the "unexplained" LFP rose for married women but not for single women. It therefore highlights, firstly, that the labor force participation 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 behavior of marrieds and that of singles 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.

2.2 Shift in work patterns upon marriage

To further corroborate the observation that married individuals' labor supply choice is disparate from singles', I study how individuals shift their time use immediately upon marriage, and how this shift changes over time.

I follow the event-study approach of Kleven and Landais (2018). 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 housework and market work 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 age 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 level of 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].

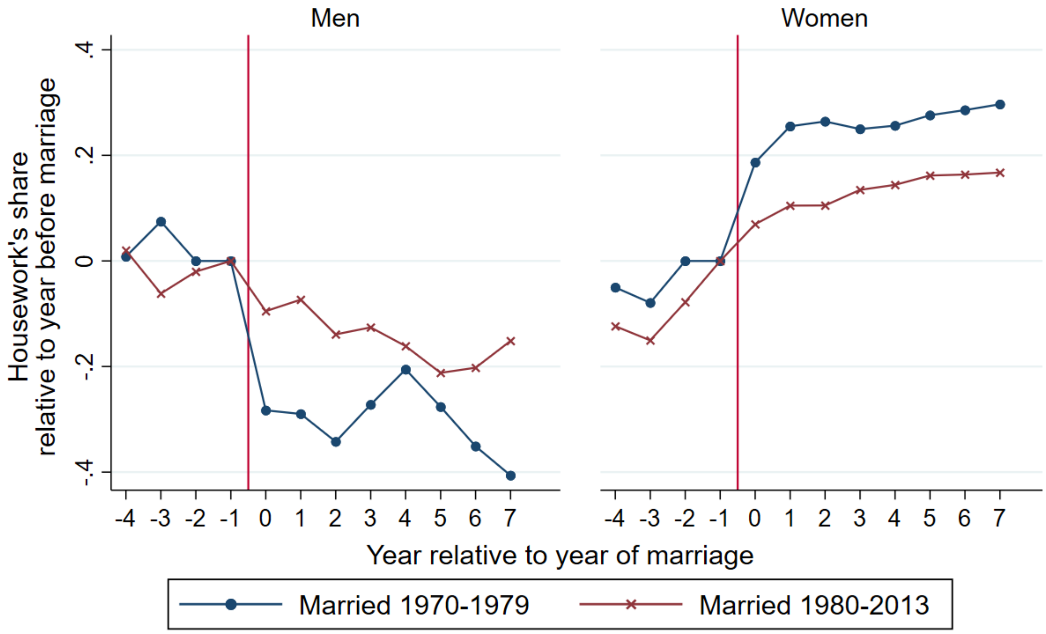
The blue lines of Figure 2 illustrate a sharp jump in women's share of housework hours among total work (market work and housework) hours and a sharp drop in men's, immediately upon transition from singlehood to marriage, for marriages in the 1970s. Thus the marrieds of the 1970s faced additional constraints that made them engage more in the activity traditionally prescribed to them; women do more housework and men more market work. The magnitude of the time use shift is also not trivial. The jump in the housework's share for women upon marriage 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 individuals to the gender role constraints. This idea resonates with the catch-up of married women's LFP with single

[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

women's even when residualized for the number of children, shown in Figure 1.

Figure 2: Housework's share of housework and market hours, among couples whose first child is born $\geq$ 4 years after marriage



Notes: This figure plots the share of housework among the sum of housework and market work 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.

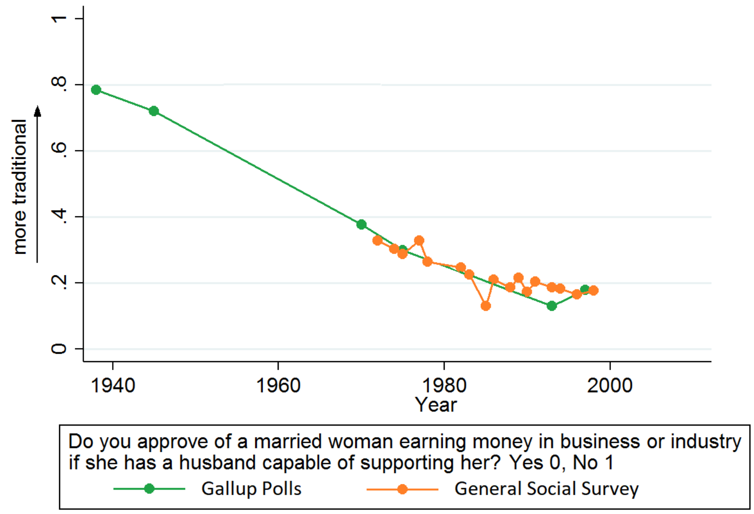
The red lines of Figure 2, on the other hand, demonstrate that there are no sharp changes in time allocation for marriages that take place later. The event study coefficients are also statistically insignificant around the year of marriage. I take this null effect for later marriages as suggestive of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

2.3 Attitudinal survey trends

The last motivating fact supports the notion of weakening gender roles over time.

Figure 3 illustrates that attitudes on the gender role of married women have become less traditional over time. Among various survey questions on gender attitudes, 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. While close to 80% answered 'No' to this question 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.

Figure 3: Trend in attitudinal survey answers



Notes: 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). Although the two surveys asked almost identical questions, there is a slight difference. The Gallup Polls' specific question was "Should a married woman earn money if she has a husband capable of supporting her?" with possible answers "Yes" (0) and "No" (1), 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?" with possible answers "Yes" (0) and "No" (1).

Other survey questions on gender roles of married individuals have been asked, however, with answers confirming a continual trend towards less traditional attitudes. These trends are shown in Figure A1 of the Appendix.

3 Model

The motivating facts of Section 2 suggest that married individuals' labor supply decisions are different from single individuals', that this gap is shrinking over time, and that gender roles could be one of the factors driving this gap.

I proceed by building a structural model, for two purposes. I first use it as a measurement tool to quantify by how much gender roles affect married individuals' labor supply choice. Then I use the model to conduct counterfactuals to gauge their importance for various aggregate measures related to labor supply, earnings, marriage, and education.

Timing In my model, the decisions of individuals are divided into three stages. In the first stage, they choose their level of education as a forward-looking investment decision, balancing the returns to education in the labor and marriage markets and the cost of education (Chiappori, Iyigun, and Weiss, 2009). In the second stage, they enter a frictionless transferable utility (TU) marriage market (Becker, 1973; Shapley and Shubik, 1971), where "types" of individuals equals 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 individuals and by one's own education level for single individuals. 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). Individuals then 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 wages that the representative firm of the economy pays, b) the group-specific value of home production, and c) the group-specific disutilities that a married couple gets upon deviation from traditional gender roles, i.e. when the wife works in the market and when the husband works at home. 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 starting from the last.

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

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 is the consumption of the private good, and Q is the consumption of the public good (e.g. expenditures on housing, children, heating). $L_f \in \{0, 1\}$ denotes the labor market participation of wife f, and w_i refers to the market wage of individual i. Critically, married individual i gets disutility when labor allocations deviate from traditional gender norms. The disutilities, parameterized by τ, occur proportionally to the value of market income brought home by the wife.

From an ordinal perspective, this utility belongs to Bergstrom and Cornes' Generalized Quasi Linear (GQL) family. Hence, at any period and for any realization of family income, it satisfies

[8] This paper formulates an equilibrium lifecycle model of education, marriage and labor supply and consumption in a transferable utility context. A key innovation of this paper is that labor supply decisions are made over the lifecycle, while maintaining equilibrium in the educational choice and marriage matching.

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,[9] any cardinalization of u can be used for the definition of joint marital output. 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, $\exp u_m + \exp u_f$.[10]

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

$$\max_{Q,C} Q(C - \tau w_f L_f) \tag{2}$$

s.t.

$$pQ + C = w_m L_m + w_f L_f + h_m(1 - L_m) + h_f(1 - L_f)$$

$C \equiv C_m + C_f$ 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, $L_m \in \{0, 1\}$ denotes the labor market participation of husband m, and h_i refers to the home productivity of individual i.[11] The solutions to the maximization problem given by (2) are

$$Q = \frac{w_m L_m + (1 - \tau) w_f L_f + h_m(1 - L_m) + h_f(1 - L_f)}{2p} \tag{3}$$

$$C = pQ + \tau w_f L_f$$

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),[12] which is endogenously determined in the marriage market:

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

---

[9]The set of Pareto efficient allocations remains unchanged when u is replaced with $f(u)$, for a strictly increasing mapping f.

[10]Proof: An allocation is Pareto efficient if it maximizes $\exp u_m$ subject to (a) the budget constraint, and (b) $\exp u_f \geq \bar{u}$. This program is equivalent to a second program that maximizes $\exp u_m + \zeta \exp u_f$ subject to the budget constraint alone, where ζ is the Lagrange multiplier on constraint (b). The first order conditions of the second program with respect to private consumptions yield

$$Q = \lambda = \zeta Q$$

where λ is the Lagrange multiplier on the budget constraint. Thus, it must be that $\zeta = 1$; any Pareto efficient allocation maximizes the sum of exponential utilities.

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

[12]See section 6.3.2 of Browning, Chiappori, and Weiss (2011) for more detail on the characterization of intra-household allocations under efficient risk-sharing.

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.

Clearly, the couple makes labor choices to maximize Q.[13] Hence, from (3), the optimal labor choices are

$$L_m^* = \mathbb{1}\left[w_m \geq h_m\right] \tag{6}$$

$$L_f^* = \mathbb{1}\left[(1-\tau)w_f \geq h_f\right] \tag{7}$$

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. Another feature that stands out in equation (6) is the independence of the husband's and wife's labor supply choices. See Appendix C.4 for a discussion on this feature.

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 do not receive disutilities from non-traditional behavior.

A single individual i maximizes the following utility:

$$\hat{u}_i(\hat{Q}_i, \hat{C}_i) = \ln(\hat{Q}_i) + \ln(\hat{C}_i) \tag{8}$$

s.t.

$$p\hat{Q}_i + \hat{C}_i = w_i\hat{L}_i + h_i(1-\hat{L}_i)$$

The resulting indirect utility is

$$\hat{v}_i = 2\ln \hat{Q}_i + \ln p \tag{9}$$

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

$$\hat{L}_i^* = \mathbb{1}\left[w_i \geq h_i\right] \tag{10}$$

[13] 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.

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. Gender, marriage match, and family composition are all determined before the labor choice stage. After the marriage matching stage, family composition is given exogenously according to match-specific empirical probabilities.[14] Hence, the probability that a (husband type q, wife type r) match has a family composition $\mathcal{K}$, denoted as $d^{qr}(\mathcal{K})$, is simply found from the data.

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 type 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 tax parameters from home productivity parameters.

I assume that market abilities ϵ^w and home production abilites ϵ^b are drawn independently and identically from the Fréchet distribution with shape parameter θ, after the marriage matching stage.[15]

[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.

[15] The extensive literature on returns to schooling 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, $\epsilon_s^w = c^s\epsilon^w$. Then, it is possible to take the scaling constants (c^s) out of intrinsic abilities and have schooling-specific wages incorporate the scaling constants.

The cumulative distribution functions for these abilities are

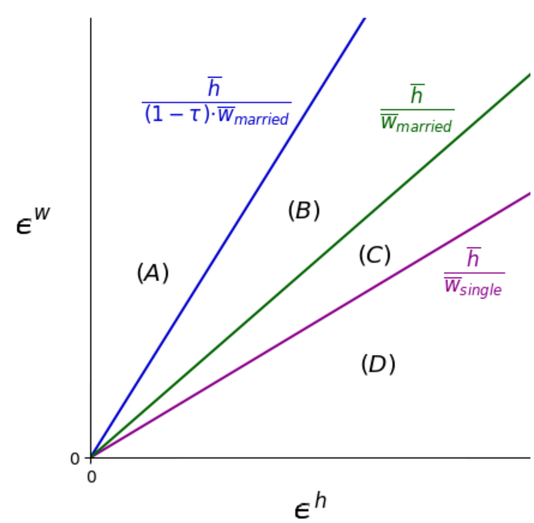
$$F(\epsilon^w) = F(\epsilon^b) = F(x) = \exp\left\{-x^{-\theta}\right\}.$$

From the convenient property of Fréchet distributions (Eaton and Kortum, 2004), the probability that a woman in a (q, r) match with family composition $\mathcal{K}$ works in the labor market is:

$$P_F^{qr}(\mathcal{K}) \equiv \mathbb{P}\left((1-\tau^{qr})\bar{w}_F^{qr}(\mathcal{K})\epsilon^w > \bar{h}_F^r(\mathcal{K})\epsilon^b\right) = \frac{[(1-\tau^{qr})\bar{w}_F^{qr}(\mathcal{K})]^\theta}{[(1-\tau^{qr})\bar{w}_F^{qr}(\mathcal{K})]^\theta + [\bar{h}_F^r(\mathcal{K})]^\theta} \tag{11}$$

The maximum likelihood estimator for this probability is the labor force participation rate of the women in this group. Equation (11) 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.

Figure 4: Sorting across market work and home production of married and single women



Notes: 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^b, 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 female norms wedge was removed, more married women would be engaging in market work.
(C): Along with (A) and (B), the market and home ability combinations of single women doing market work.
(D): Single women who work at home are very talented in home production.

Another implication of Fréchet abilities useful for calibration later on is that the average wage of the women working in the market is[16] :

$$\begin{aligned} \text{avgwage}_F^{qr}(\mathcal{K}) &= \bar{w}_F^{qr}(\mathcal{K})\mathbb{E}\left[\epsilon^w|(1-\tau^{qr})\bar{w}_F^{qr}(\mathcal{K})\epsilon^w > \bar{h}_F^r(\mathcal{K})\epsilon^b\right] \\ &= \bar{w}_F^{qr}(\mathcal{K})\left(\frac{1}{P_F^{qr}(\mathcal{K})}\right)^{\frac{1}{\theta}}\Gamma\left(1-\frac{1}{\theta}\right) \end{aligned} \tag{12}$$

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:

$$n^{q0} + \sum_{r=1}^{S} n^{qr} = M^q \qquad \forall\, q = 1, ..., S \tag{13}$$

$$n^{0r} + \sum_{q=1}^{S} n^{qr} = F^r \qquad \forall\, r = 1, ..., S \tag{14}$$

In this TU model, a type-q man must transfer τ^{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

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

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

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

where ε_f^{qr} is f's random preference for the match drawn independently and identically from the type

[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 $\left(\frac{1}{P}\right)^{\frac{1}{\theta}}$.

I extreme-value distribution.

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

$$n^{qr,\mathcal{D}} = n^{qr,\mathcal{S}} = n^{qr}.^{17}$$

In equilibrium,[18]

$$\frac{n^{qr}}{\sqrt{n^{q0}n^{0r}}} = \frac{\mathbb{E}(v_m^{qr}) + \mathbb{E}(v_f^{qr}) - \mathbb{E}(\hat{v}_m^q) - \mathbb{E}(\hat{v}_f^r)}{2} + \psi^{qr}$$

Using equation (5),

$$\frac{n^{qr}}{\sqrt{n^{q0}n^{0r}}} = \frac{\mathbb{E}(\breve{v}^{qr}) - \mathbb{E}(\hat{v}_m^q) - \mathbb{E}(\hat{v}_f^r)}{2} + \Psi^{qr} \tag{15}$$

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 (15) is the gain to 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.[19]

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:

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

where

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

Individuals are forward-looking. The expected utility from schooling level r depends on the consequent matching probabilities in the marriage market and the expected utilities in each type of match. The costs, 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.

[17] There is a 1:1 relationship between τ^{qr} and the Pareto weight μ^{qr}. From the marriage market clearing condition and equation (4),

$$\tau^{qr} = \frac{\ln n^{q0} - \ln n^{0r} - \ln \mu^{qr} - \mathbb{E}(\hat{v}_m^q) + \mathbb{E}(\hat{v}_f^r)}{2}$$

[18] See Appendix section C.5 for greater details on the derivation of the marriage market equilibrium.

[19] ψ^{qr} and μ^{qr} cannot be separately identified. As can be seen in footnote 17, 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, conditional on the behavioral assumption of limited foresight, described in section 4.2.

From equation (4),

$$U_F^r = 2\sum_{q=0}^{S}\left[\frac{n^{qr}}{F^r}\mathbb{E}(\ln Q^{qr})\right] + \ln p - C_F^r - \xi^r \tag{16}$$

where $C_F^r \equiv c_F^r - \sum_{q=0}^{S}\frac{n^{qr}}{F^r}\left(\ln\frac{\mu^{qr}}{1+\mu^{qr}} + \tau^{qr} + \psi^{qr}\right)$. C_F^r is the direct cost of getting schooling level r, minus a) the expected utility from intra-household resource allocation, b) the marriage market utility transfer, and c) the noneconomic gains in a match. In the parameter inference section (section 4), I will back out the values of C_F^r, and not c_F^r.[20]

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 $\mathcal{F}$:

$$Y^{mkt} = B(\mathcal{M} + \mathcal{F}) \tag{17}$$

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. $\bar{w}_M^{10}(\mathcal{K}_0)$.

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

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

[20] As with Ψ^{qr}, identifying C_F^r is sufficient for running counterfactuals, conditional on the behavioral assumption of limited foresight, described in section 4.2.

3.5 Aggregate Output

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

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

where

$$Y^{home} = B \sum_{g \in \{M,F\}} \sum_{(q,r)} \sum_{\mathcal{K}} n^{qr} d^{qr}(\mathcal{K}) \left(\frac{\bar{h}_g^r(\mathcal{K})}{\bar{w}_M^{10}(\mathcal{K}_0)} \right) \left(1 - P_g^{qr}(\mathcal{K})\right)^{1-\frac{1}{\theta}} \Gamma \left(1 - \frac{1}{\theta}\right) \tag{18}$$

3.6 Equilibrium

An equilibrium in this economy consists of schooling choice q for a man, schooling choice r for a woman, marital transfers τ^{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 $\{\tau^{qr}\}$ 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 (8) 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 (17).
5. Market output Y^{mkt} is given by equation (17), and total home production Y^{home} by (18).
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 the non-traditional working arrangement when married is lower. Then some of the 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 the 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, young women realize that they are more likely to be married and to work in the labor market in the future. Then if there is positive assortative matching on education in the marriage market, the greater likelihood of marriage increases their incentive for higher education. This effect comprises the marriage-market returns to education. In addition, if education is more effective in increasing market productivity than home productivity, the greater likelihood of market work increases young women's incentive for higher education. This effect, on the other hand, comprises the labor market returns to education. Either case, higher education for women would then increase aggregate output 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.[21]

The model in section 3 is fitted to the census data every decade, assuming that the data is a reflection of the model steady state. By calibrating the model separately by decade, I am allowing almost 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 $\{\bar{h}_M^{qr}(\mathcal{K}), \bar{h}_F^{r}(\mathcal{K})\}$, 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 only parameter that I leave to be constant over time is θ, the inverse measure of the dispersion of market and home abilities.[22]

[21] Appendix figure A5 shows that the age range of 25-54 is appropriate as the most economically active 30-year window. Hsieh et al. (2019) also use this age range.

[22] See section 4.4 for more discussion and for results when θ is estimated for each decade.

The practical advantage of my model set-up is that the model is not demanding on the data; the only 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 spousal information.

4.2 Assumption on behavior under counterfactual scenarios

Identifying parameter values is necessary to conduct counterfactual simulations. Under counterfactual situations, the marriage matching pattern will be different and the match-specific Pareto weights may change as a result. However, it is impossible to figure out what the counterfactual Pareto weights would be, without imposing further structure on how they are determined. Therefore, I assume that individuals are naive with limited foresight; they expect future marriage market outcomes[23] to remain the same under counterfactual scenarios. This assumption makes identifying Ψ^{qr} and C_g^s, rather than separately identifying μ^{qr}, ψ^{qr} and c_g^s, sufficient for deriving counterfactual marriage and education patterns.

4.3 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:[24]

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

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 (12).

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

[23] Specifically, the exact objects that individuals need to expect unchanged under counterfactual scenarios are the a) Pareto weights $\{\mu^{qr}\}$, b) probabilities of matches $\{\frac{n^{qr}}{F^r}, \frac{n^{qr}}{M^q}\}$, and c) marriage market equilibrium transfers $\{\tau^{qr}\}$.

[24] 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. $\bar{h}_g^{qr}(\mathcal{K})$: group home productivity per unit of effective labor

   $\bar{h}_F^r(\mathcal{K})$ (similarly, $\bar{h}_M^{qr}(\mathcal{K})$) can be backed out from equation (11), 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 $\mathcal{K}$.

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

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

   $\tau^{qr}(\mathcal{K})$ is backed out from equation (11) using $\bar{w}_F^{qr}(\mathcal{K})$, $\bar{h}_F^r(\mathcal{K})$, 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.

$$\tau^{qr}(\mathcal{K}) = 1 - \frac{\text{avgwage}_F^{0r}(\mathcal{K})}{\text{avgwage}_F^{qr}(\mathcal{K})}\left(\frac{1 - P_F^{0r}(\mathcal{K})}{1 - P_F^{qr}(\mathcal{K})}\right)^{\frac{1}{\theta}} \tag{19}$$

   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 $\mathcal{K}$. 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 (15),

$$\Psi^{qr} = \frac{n^{qr}}{\sqrt{n^{q0}n^{0r}}} - 2A^{qr} + \hat{A}_M^q + \hat{A}_F^r. \tag{20}$$

   where

$$A^{qr} = \sum_{\mathcal{K}} d^{qr}(\mathcal{K})\mathbb{E}\Big[\ln\Big(\bar{w}_M^{qr}(\mathcal{K})\epsilon_m^w L_m^* + \bar{h}_M^{qr}(\mathcal{K})\epsilon_m^h(1-L_m^*) + \big[1-\tau_F^{qr}(\mathcal{K})\big]\bar{w}_F^{qr}(\mathcal{K})\epsilon_f^w L_f^* + \bar{h}_F^r(\mathcal{K})\epsilon_f^h(1-L_f^*)\Big)\Big]$$

   and

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

   There are no closed-form expressions for A^{qr} and $\hat{A}_g^s$ so I simulate them to back out Ψ^{qr}.[25]

7. C_g^s: direct cost minus expected utility from intra-household resource allocation, of education s

[25] Given all the parameter values found in steps 1-5, the simulation is straightforward.

Use equation (16). $\{C_F^r\}_{r=1,...,S}$ are found as the solution to the system of equations

$$\frac{F^r}{\sum_{r'=1}^{S} F^{r'}} = \frac{\exp\left\{2\sum_{q=0}^{S}\frac{n^{qr}}{F^r}A^{qr} - C_F^r\right\}}{\sum_{r'=1}^{S}\exp\left\{2\sum_{q=0}^{S}\frac{n^{qr'}}{F^{r'}}A^{qr'} - C_F^{r'}\right\}} \qquad \forall\, r = 1, ..., S$$

where $A^{0r} = \hat{A}_F^r$ and $A^{q0} = \hat{A}_M^q$. $\{C_M^q\}_{q=1,...,S}$ are found similarly.

4.4 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 estimate of θ is 1.837, which is similar to Hsieh et al. (2019)'s estimate of 1.52 for the Fréchet shape parameter dictating the dispersion of abilities across occupations. It is also close to their choice to use 2 for conducting counterfactuals. θ is estimated for the entire sample from 1940 to 2010, under the assumption that the distribution market and home ability endowments remains fixed over time. If θ is estimated decade by decade, the estimates are quite stable although there is a slight non-monotonic upward trend. They range from 1.549 to 1.999, with 1.844 as the median value.[26] 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.

Table 1: Maximum likelihood estimate of θ

$\hat{\theta}$	1.837***
	(18.31)
N	3570573

Notes: t statistics based on 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 at 4579.11 due to the large sample size.
See step 2 of section 4.3 for the maximum likelihood estimation strategy.

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

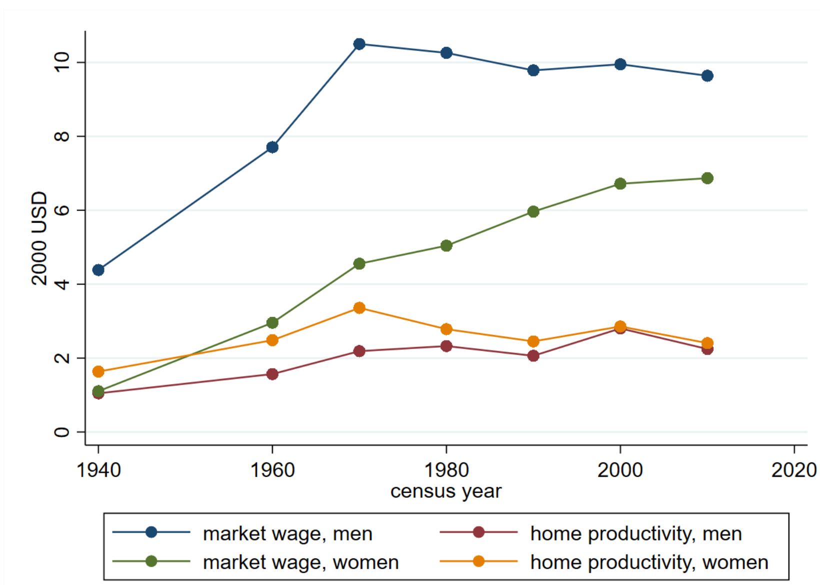
To calibrate the group-specific market and home productivities, I first need to specify the groups. Each group is defined by gender, schooling pair, and family composition. I must ensure that each

[26] The θ estimates by decade are 1.549 in 1940, 1.844 in 1960, 1.845 in 1970, 1.718 in 1980, 1.975 in 1990, 1.999 in 2000, and 1.826 in 2010.

group is large enough since I match population moments to sample analogs within each group.[27] 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. 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. This categorization is in Table A2 of the Appendix. A similar rationale holds for family composition categories; I construct these 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. The family composition categories, detailed in Table A3 of the Appendix, are kept fixed over the entire period.

Figure 5 plots the weighted average of group market wage and home productivity by sex and decade. 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 the computation of 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



Notes: 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 4.3.

The group market productivity $\bar{w}$ is increasing in the average wage of the workers in that group

[27]For example, see equations (11) and (12), applied in the parameter inference steps 3 and 4 of section 4.3.

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 effects. 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.

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. 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 the early decades of 1940 and 1960. Thus for home productivity, the effect of education on group LFP rate often dominates that on the average wages of the group's workers. This finding contrasts with the literature on the positive returns to education on childcare (Leibowitz, 1974). It could also 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[28] 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

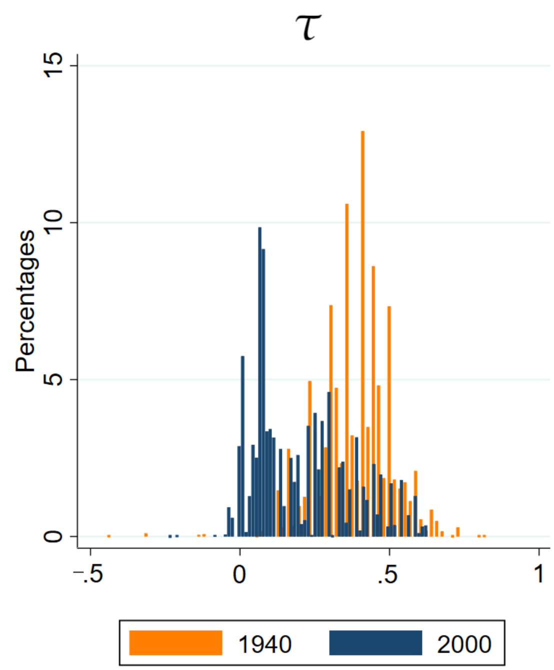
I next calibrate the values of τ, the norms wedges on the market wage of married women. Though labeled as "norms wedges", τ encapsulate any reason that brings married women's LFP to diverge from single women's LFP besides wage differentials. For instance, it includes the differential valuation of staying home between marrieds and singles.

To illustrate the interpretation of τ_F and τ_M through the lens of the model, let us take an example. If τ_F equals 0.6, the interpretation is that the worth of a \$10 market wage to a married woman is only \$4 when she is making her labor supply decision. Thus, $\tau_F = 0.6$ corresponds to a norms wedge of 60% on the market wage of that woman. Similarly, $\tau_M = 0.3$ corresponds to a norms wedge of 30% on the value of home production of a married man.

[28] e.g. Ngai and Petrongolo (2017)

Figure 6 plots the histograms of τ_F and τ_M, 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 histograms of both τ_F and τ_M for 2000 sit to the left of those for 1940, signifying a decrease in the norms wedges. Moreover, the distribution of τ_M is more dispersed than that of τ_F, because of the nonlinearity in the closed-form expressions for τ_F and τ_M in equations (19) and (**??**).

Figure 6: Histogram of norms wedges on married women's market wages (τ)



Notes: 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 4.3.

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.[29] Figure 7 plots these values. τ generally decreases over time, except from 2000 to 2010. The reason for this rise is that while married women's wages have increased relative to single women's between the two decades, their LFP has not.[30,31] To reconcile these two observations, τ must increase since τ encapsulates any reason for which the LFP rates of married women and single women diverge besides wage differentials. A possible interpretation is that the value of home production of married women has increased.[32]

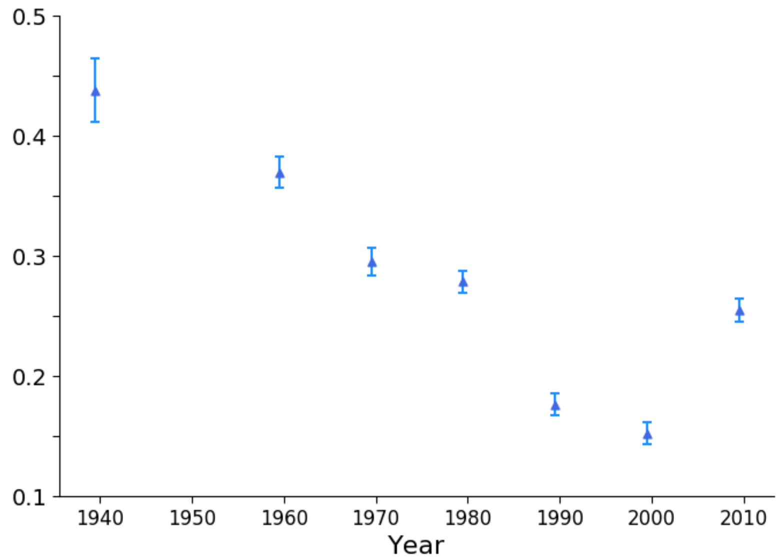
[29] The results are very similar to weighted average, unweighted average, or unweighted median values.

[30] Figure 1 illustrates that female labor force participation has plateaued since 1990 regardless of marital status.

[31] Even when I take into account the market work *hours* of married women relative to single women, the rise persists.

[32] The IPUMS Time Use survey indicates a rise in child care hours of married women relative to single women recently.

Figure 7: Evolution of inverse norms wedges τ_F and τ_M



Notes: This figure plots the weighted median of τ, 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.

To show that the values of norms wedges imputed are related to directly observable measures of conservativeness, 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.[33] Then I take for each state either the weighted average or the weighted median of the norms wedges τ. Table 2 reports the regression coefficients from regressing the state-level norms wedges on the state-level attitudinal survey answers. The state-level attitudinal survey answers are the weighted average by state of individual survey answers taken across multiple periods. Two attitudinal measures are considered. The first is the fraction disapproving of married women working, using the data on 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.

[33]The age cohorts are ages 25-34, 35-44, and 45-54.

Table 2: Correlation between state-level norms wedge and attitudinal survey answers

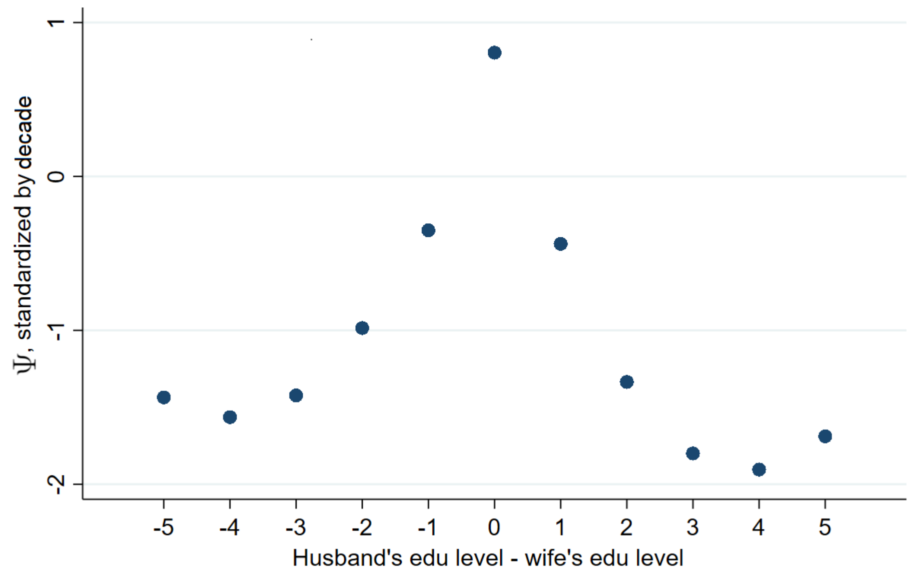
	Dependent variable τ	
	average	median
Regressed on:		
Fraction disapproving of married women working	0.249**	0.282**
	(2.21)	(2.21)
Regressed on:		
Composite attitudinal index	0.450***	0.439**
	(2.94)	(2.50)
N	51	51

Notes: t statistics based on 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_F 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_F 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).

Figure 8: Ψ by spousal education gap



Notes: 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: direct cost minus expected utility from intra-household resource allocation

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, are not comparable across time since the schooling level categories differ by decade.[34] However, the costs can be compared between the two genders within each decade. The cost of attaining the highest schooling level was larger for women from 1940 to1990, but it became *smaller* for women from 2000. 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. The share of women of the same age range with at least some graduate school overtook that of men around 2000, too. Furthermore, a time-series comparison can be made for the decades 1990-2010, as the schooling categories are the same for those decades. The female to male ratio of the cost of the highest schooling level steadily falls from 1.11 in 1990 to 0.96 in 2000, and to 0.94 in 2010.

[34] Refer to Table A2 in the Appendix for the schooling categories.

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 (20) as well as the $2S$ accounting identities given by (13) and (14).

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 2010 if gender norms had remained at the level of 1940, holding all other parameter values fixed at the 2010 level. As recognized in section 4.4, the calibrated norms wedges for *each group* are most likely noisy estimates. For this reason, I consider the counterfactual of every individual in 2010 being subject to the same, weighted-median norms wedge of 2010 at baseline, and being subject to the same, weighted-median norms wedge of 1940 in the counterfactual scenario.

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 0.25 in 2000 to 0.44 in 1940, i.e. more traditional gender norms for married women, married women work 14.3% less in the labor market. As a result, the cumulative market output Y^{mkt} of married women falls by 6.9%. However, as fewer married women work in the market, married women's cumulative home production value Y^{home} increases, and so the total output $Y(= Y^{mkt} + Y^{home})$ falls by less, at 2.1%. 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 female norms wedge of 1940

	Adjustment margins	
	Labor supply (1)	Education, marriage, & labor supply (2)
Education		
Women's years of schooling	-	-1.4
Men's years of schooling	-	-0.8
Selection into marriage		
Marriage rate	-	-32.2
Married women's edu/single women's edu	-	-4.1
Married men's edu/single men's edu	-	-1.2
Labor Force Participation		
Married women's LFP	-14.3	-17.5
Married men's LFP	-	-0.03
Single women's LFP	-	0.6
Single men's LFP	-	0.1
Output per head		
Married women's market output	-7.0	-13.0
Married women's total output	-2.1	-6.5
Married men's market output	-	-0.8
Married men's total output	-	-0.8
Aggregate market output	-2.0	-4.8
Aggregate market & home output	-0.6	-3.5
Within-household gender earnings gap		
Wife's share of household market income	-11.5	-14.9

Notes: This table reports the percentage changes in various aggregate variables that occur when the individuals of 2000 are subject to the female 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 32.2%. As the norms wedge is more costly for women with higher market ability, the fall in marriage is more pronounced for higher educated women. Because there is assortative marriage matching by education, the fall in marriage is more pronounced also for higher educated *men*. Furthermore, the education of both men and women fall.[35] With a higher τ, the labor and marriage market returns to education are lower for women, and only the marriage market returns to education are lower for men. As the set of married women has a lower market ability on average with higher τ, 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 6.5%. In aggregate, including the output of single men and women, aggregate market output falls by 4.8% and total output by 3.5%.

In summary, the fall in gender norms wedges over 1940 through 2010 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) faster growth of educational attainment of women relative to men, and d) increasingly positive selection of men and women into marriage by education (Bar et al., 2018; Juhn and McCue, 2016; Case and Deaton, 2017).

Is the effect of a 4.8% fall in aggregate market output and a 3.5% 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. On the other hand, it might be viewed as small, relative to the output growth that has occurred over 1940-2010. To better benchmark the size of the 2% effect, I conduct additional counterfactuals.

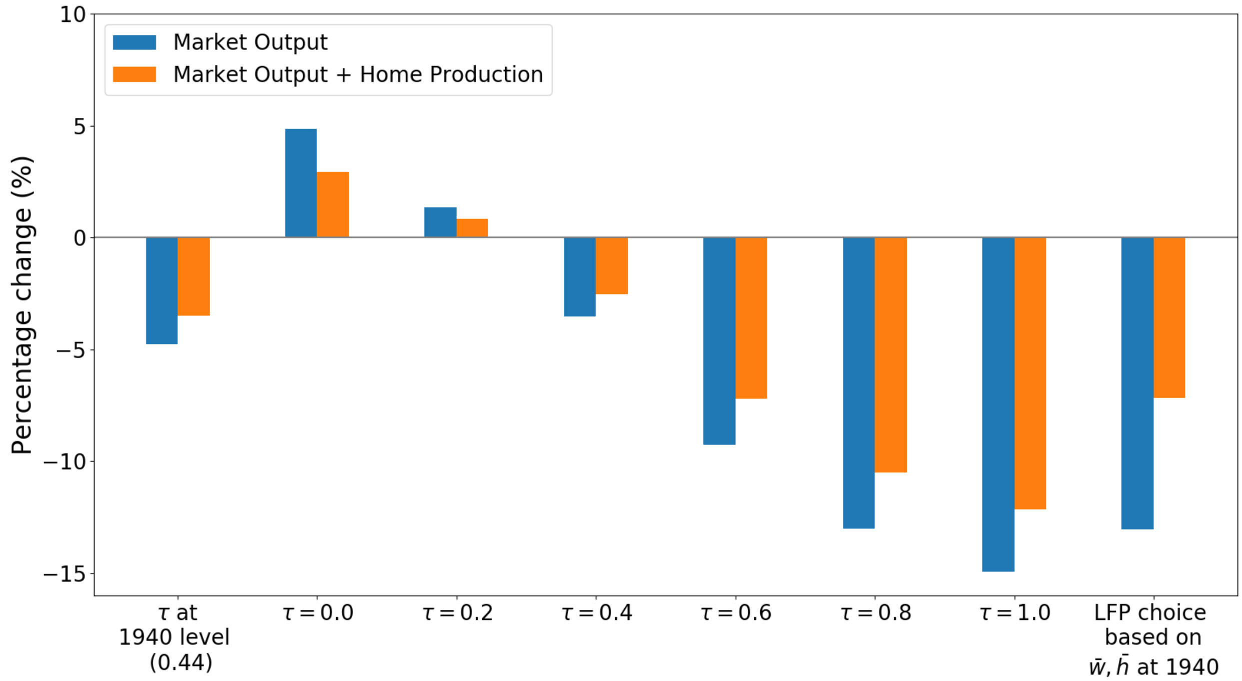
Additional Counterfactuals Figure 9 compares the market and total (market and nonmarket) output effects of various counterfactual scenarios, where the baseline year is set at 2010. The other counterfactual scenarios explore the effects of τ changing from 0 to 1, and when labor supply choices are made based on $\bar{w}$ and $\bar{h}$ of 1940. All the three margins of education, marriage, and labor supply are allowed to adjust.

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, underlying in great part the output growth over 1940-2010. In fact, $\bar{w}$ is lower than $\bar{h}$ for women in 1940, whereas $\bar{w}$ is *more than triple* of $\bar{h}$ in 2010. Not to include the direct effects of productivity changes on output, the counterfactual is only about letting labor supply choices be determined based on $\bar{w}$ and $\bar{h}$ at 1940 levels. Then, the labor supply choice changes enormously, with married women's LFP rate falling by a staggering 66%. Market output consequently falls by 13.1% and total output falls by 7.2%. Therefore, the main counterfactual's effects amount to a half of the

[35] In an attempt to succinctly express the change in education, I compute for each decade, the weighted average of schooling years of individuals in every education level. The baseline and counterfactual compare the sum of (weighted average of schooling years by education level)×(share of population in each education level). Although the degree of the decline in the years of schooling appears very small, a lot changes with the education *levels*.

effect of the last counterfactual. In this sense, the effect of the norms wedge is sizable.

Figure 9: Market and total output effects of various counterfactual scenarios



Notes: This figure compares the market and total (market and nonmarket) output effects of various counterfactual scenarios, where the baseline year is set at 2010, with τ at 2010 being 0.25.

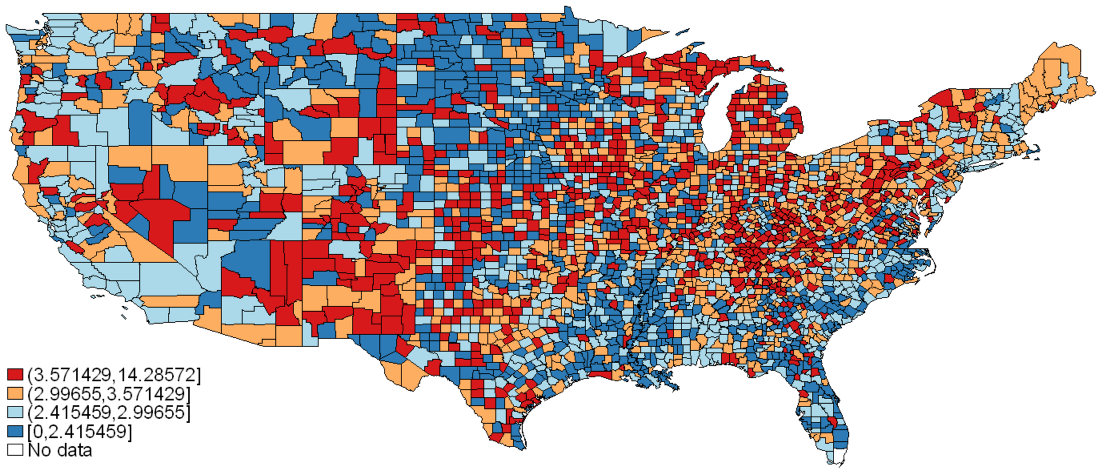
6 Reduced Form Exercise

The counterfactual results in the previous section depend on the model structure. The extended discussion on the calibrated parameter values in section 4.4 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. This model extension is compatible with how I identified norms wedges previously, as long as individuals take norms as given and do not internalize the effect of their labor supply choice on the norms of future generations.

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. Inspired by Fernández et al. (2004),[36] 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) or social learning (Fernández, 2013; Fogli and Veldkamp, 2011).

Figure 10: Map of county-level draftee casualty rates



Notes: 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.

For the reduced form exercise, I match the U.S. decennial census, by county, with the WW2 military casualty records from Ferrara (2019). 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, casualties are likely to be more random than mobilization rates.

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,

$$Y_{ict} = \alpha_c + \lambda_t + \sum_{t \neq 1940} \beta_t \times casualty_c + X_{ict}\gamma + \varepsilon_{ict} \tag{21}$$

where Y represents various outcome variables, *casualty* is the county-level draftee casualty rate, and X captures pre-determined 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

[36] 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.

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 (21) 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., 2019) 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 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.[37] 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 gender norms. 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. I therefore think of a change in the *female* norms wedge as the basis of the long-term effect of WW2 draftee casualties. 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* underlying the rise in female labor force participation.

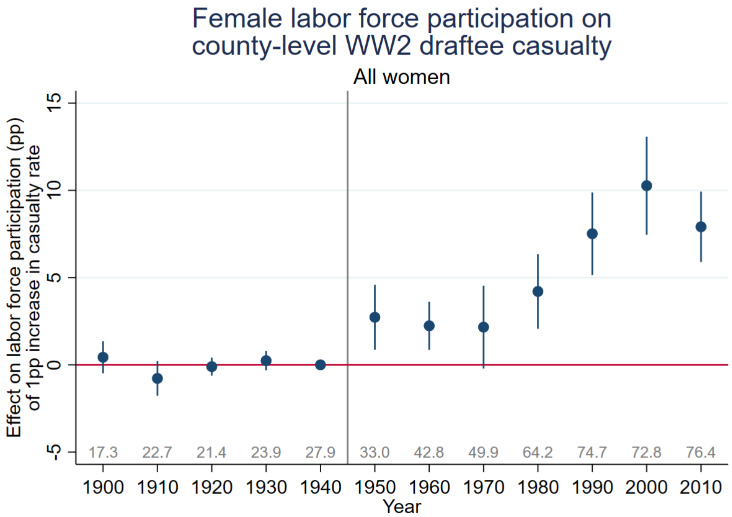
[37] There are a few caveats for plot (C). As the finest geographic variable in the attitudinal survey data is state, the difference-in-differences analysis is performed using state-level draftee casualties. Moreover, the surveys are grouped into five-year intervals. 1945 is counted as post-WW2, since 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.

The effect on female wages thus supports the model assumption of a decrease in female 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).

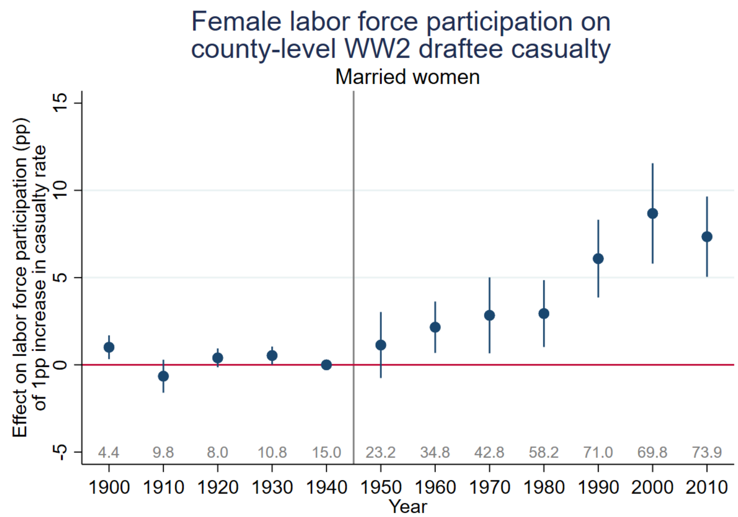
In terms of marriage and education, plot (G) depicts a rise in ever-marriage rates and plot (H) a rise in the education of women overall. Plot (H) uses whether one graduated high school or more as the measure of education, as higher levels of education are very rare to find in 1940, the pre-WW2 benchmark decade. The effect on marriage is consistent with the model assumption where norms wedges are modeled as costs to marriage, and as the female 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 female 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 female 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.

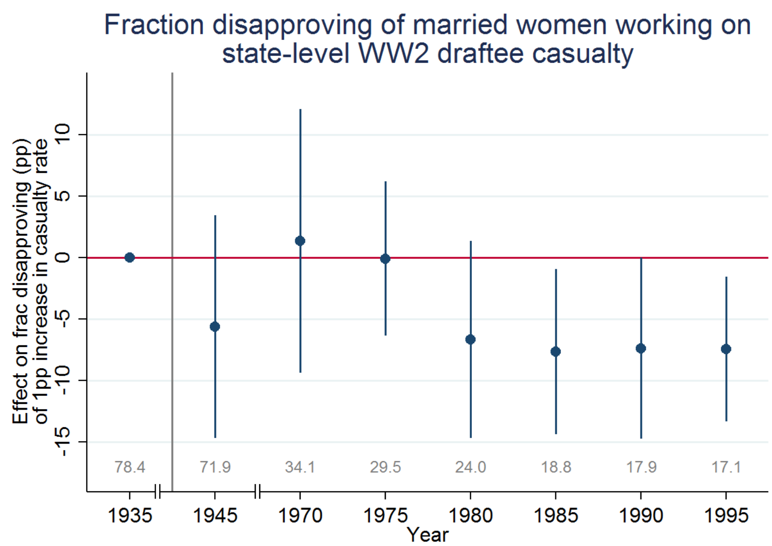
Figure 11: The effect of WW2 draftee casualty rates on various outcomes



(A)



(B)



(C)

Notes: This figure plots the difference-in-differences coefficients from estimating equation (21) 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.

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 experience 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 columns (3) and (4) of Appendix table A6. Although the coefficient sizes get smaller, both columns still demonstrate a gradual rise in female labor force participation over time.

Lastly, to further allay concerns over *level* differences in various outcome variables during the pre-WW2 period affecting the future trajectory of those variables, which would bias the usual difference-in-differences coefficients, I apply the synthetic difference-in-differences methodology of Arkhangelsky et al. (2019). Synthetic difference-in-differences estimates weights on control counties and on time periods such that the pre-WW2 path of the doubly-weighted average lies extremely close to the pre-WW2 path of the treatment counties. Hence, any level differences in the pre-WW2 period between treatment and control counties, for any outcome variable, are practically eliminated. Because of the need to divide treatment and control counties, I can only employ a binary specification. Column (5) of Appendix table A6 shows that the synthetic difference-in-differences coefficients also depict a gradual rise in female labor force participation over time.

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. Assuming that the model is in steady state each decade, there is 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 female norms wedge.

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

$$\begin{aligned} \underline{\Delta}\tau_t &= f(\underline{\Delta}FLFP_{t-1}, FLFP_{t-1}) + \nu_t \\ \approx \quad & \alpha_0 + \alpha_1\underline{\Delta}FLFP_{t-1} + \alpha_2 FLFP_{t-1} + \alpha_3\underline{\Delta}FLFP_{t-1} \cdot FLFP_{t-1} + \nu_t \end{aligned}$$

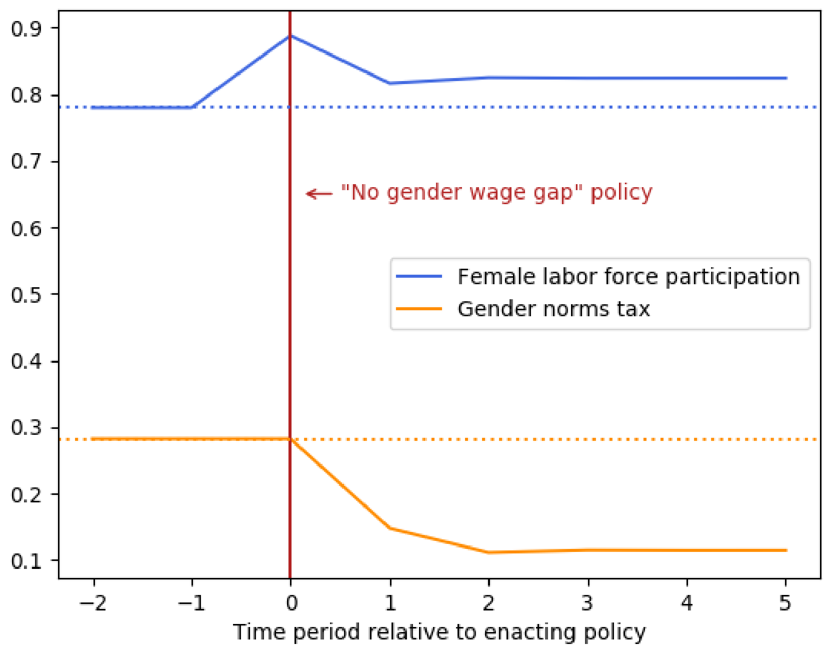
where $\underline{\Delta}$ denotes the gap between treatment and control counties. I also need two additional assumptions on how the long-term effects rise about: a) WW2 draftee casualties affect female labor force participation in 1950 and nothing else, and b) the effect only propagates through a change in the female norms wedge. These assumptions allows me to estimate the relationship between the female norms wedge and the past decade's female labor force participation that would generate the pattern of coefficients in plot (A):

$$\min_{\alpha_0,\alpha_1,\alpha_2,\alpha_3} \sum_t (\text{DID coeff, FLFP}_t - \text{change in FLFP}_t \text{ in model due to } \underline{\Delta}\tau_t)^2$$

The result is $\hat{\alpha}_0 = \underset{(0.126)}{-0.102}, \;\; \hat{\alpha}_1 = \underset{(0.185)}{0.368}, \;\; \hat{\alpha}_2 = \underset{(5.892)}{0.242}, \;\; \hat{\alpha}_3 = \underset{(7.199)}{-1.209}.$

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 steady state in a given decade in section 5, I can explore how a shock affects the model steady state *over time*.

Figure 12: The effect of paying women male wages, one-off, in 2010



Notes: This figure plots the difference-in-differences coefficients from estimating equation (21) 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.

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 is illustrative of how a one-off policy can move an economy into 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 female norms wedge to fall a decade 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 female norms wedge. Consequently, in the *following* decade, the female 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 female norms wedge than the baseline.

7 Conclusions

In this paper I measure and study the effects of gender roles associated with marriage on aggregate output, using historical data from the U.S. Gender norms became less constraining on married individuals' labor supply choices. Through direct effects on labor supply choices becoming more aligned with productivity maximization, and through indirect effects on higher education, weaker gender norms increase aggregate market and total output. Moreover, a one-off policy inducing a large rise in female labor force participation can bring an economy to a new equilibrium with higher female labor force participation.

We do not learn about development and growth only from developing countries. Rather, we can also learn from a currently developed country that has undergone large historic changes. In fact, the trend in the U.S. over the last century of gender attitudes becoming less traditional and married women's labor force participation catching up with single women's is resonated in numerous parts of the world. At the same time, one in ten countries of the world still has lower female labor force participation than 1940 U.S. (International Labor Organization, 2019). Thus, this paper can be informative 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.

A natural extension to the current paper is to take advantage of the fact that parameter identification is straightforward in this model, and apply the model to other countries. There are also 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, 2007). Building these factors into the model will allow a richer understanding of the effects of gender norms in marriage.

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 (2007). 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 (2016a). Gender roles and medical progress. *Journal of Political Economy 124*(3), 650 – 695.

Albanesi, S. and C. Olivetti (2016b). Gender Roles and Medical Progress. *Journal of Political Economy 124*(3), 650–695.

Arkhangelsky, D., S. Athey, D. A. Hirshberg, G. W. Imbens, and S. Wager (2019). Synthetic Difference In Differences. Working Paper 25532, National Bureau of Economic Research.

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 (2018). Gender Norms and the Motherhood Penalty: Experimental Evidence from India. IZA Discussion Papers 11360, Institute of Labor Economics (IZA).

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.

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

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

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 (2017). Efficiency versus Gender Roles and Stereotypes: An Experiment in Domestic Production. THEMA Working Papers 2017-10, THEMA (THéorie Economique, Modélisation et Applications), Université de Cergy-Pontoise.

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.

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

Erosa, A., L. Fuster, G. Kambourov, and R. Rogerson (2017). Hours, occupations, and gender differences in labor market outcomes. Working Paper 23636, National Bureau of Economic Research.

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.

Fernández, R. and J. C. Wong (2017). Free to leave? a welfare analysis of divorce regimes. *American Economic Journal: Macroeconomics 9*(3), 72–115.

Ferrara, A. (2019). World War II and African American Socioeconomic Progress. Working paper.

Field, E. M., R. Pande, N. Rigol, S. G. Schaner, and C. T. Moore (2019). On Her Own Account: How Strengthening Women's Financial Control Affects Labor Supply and Gender Norms. Working Paper 26294, National Bureau of Economic Research.

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

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

Goldin, C. and L. F. Katz (2002b). 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.

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.

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.

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

Juhn, C. and K. McCue (2016). Selection and Specialization in the Evolution of Marriage Earnings Gaps. *The Russell Sage Journal of the Social Sciences [Working Paper version] 2*(4).

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.

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, 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.

Leibowitz, A. (1974). Education and Home Production. *The American Economic Review 64*(2), 243–250.

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

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

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.

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 (2007). Policy distortions and aggregate productivity with heterogeneous plants. Working Paper 13018, National Bureau of Economic Research.

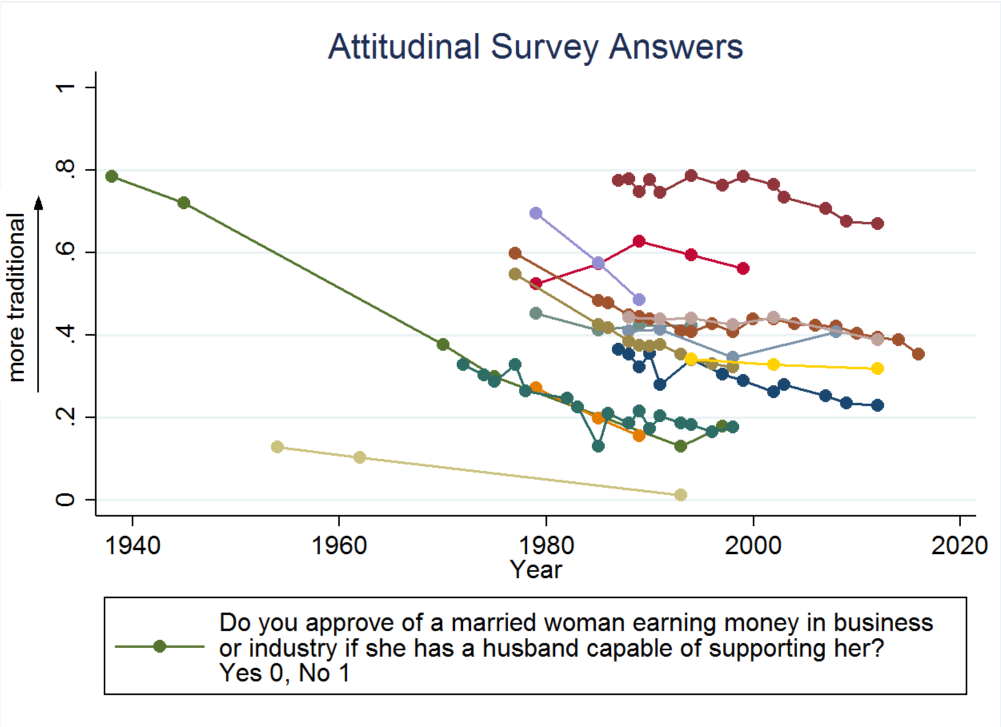
Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers 3*(2), 135–146.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2019). IPUMS USA: Version 9.0 [dataset]. Minneapolis, MN: IPUMS, 2019. `https://doi.org/10.18128/D010.V9.0`.

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

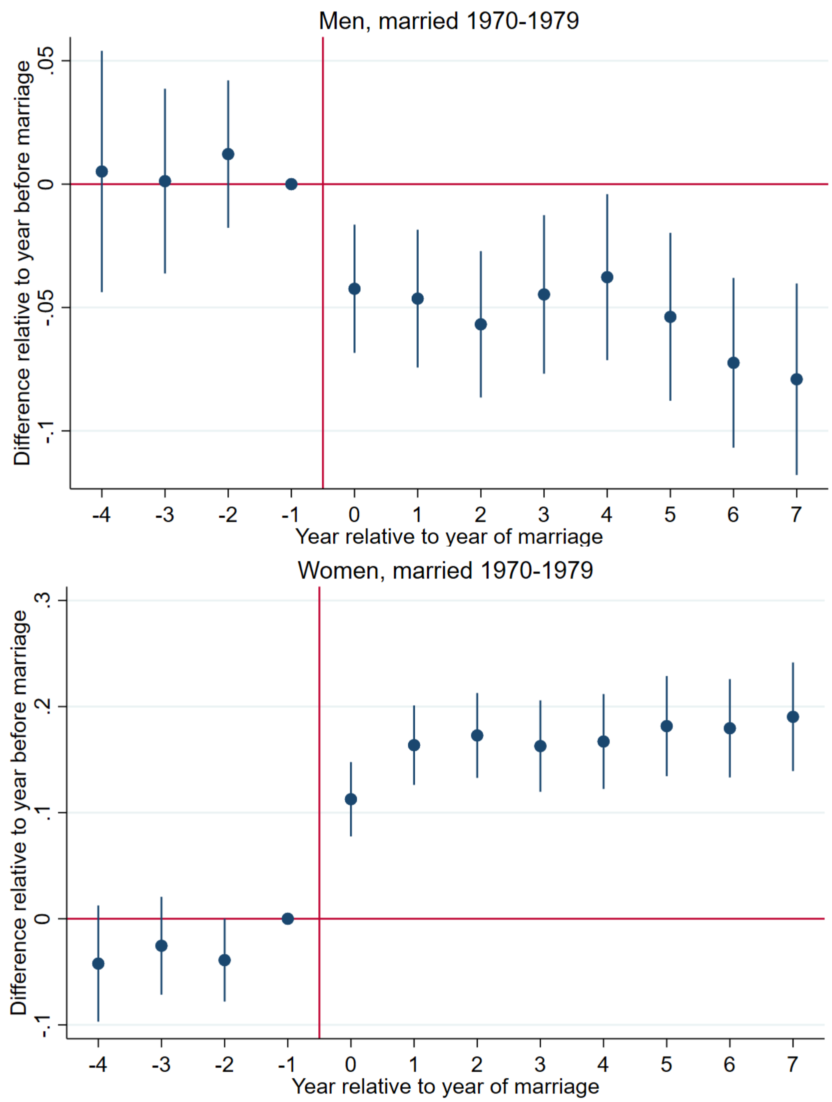
Appendix A Figures

Figure A1: Trends of answers to various attitudinal survey questions relating to gender roles within marriage



Notes: Various survey questions asked multiple times.

Figure A2: Housework's share of housework and market hours, among couples whose first child is born $\geq$ 4 years after marriage

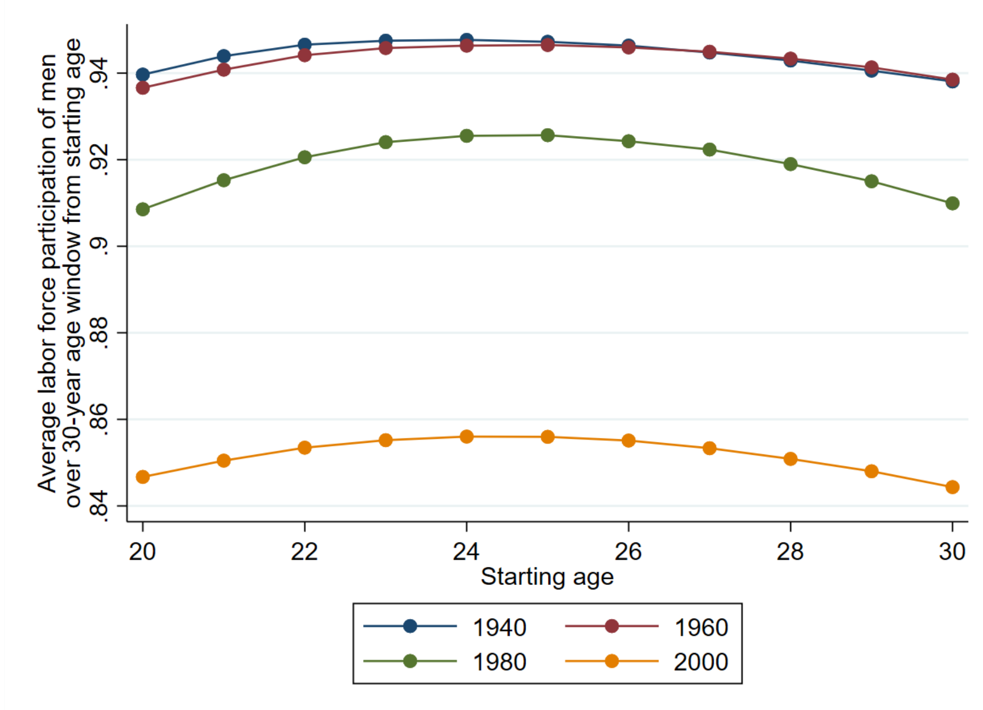


Notes: This figure plots the event-time coefficients (α_j^g) of the regression

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

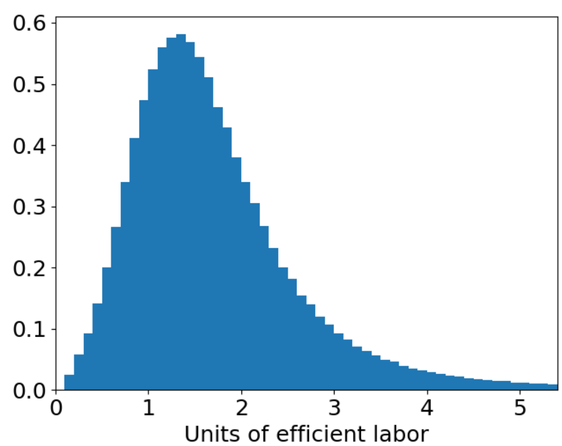
where $housework_{ist}^g$ denotes the housework's share of housework and market work hours of individual i of gender g in year s at event time t. The red vertical line plots the timing of marriage. Individuals are unmarried household heads without any live-in partners in the years to the left of the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



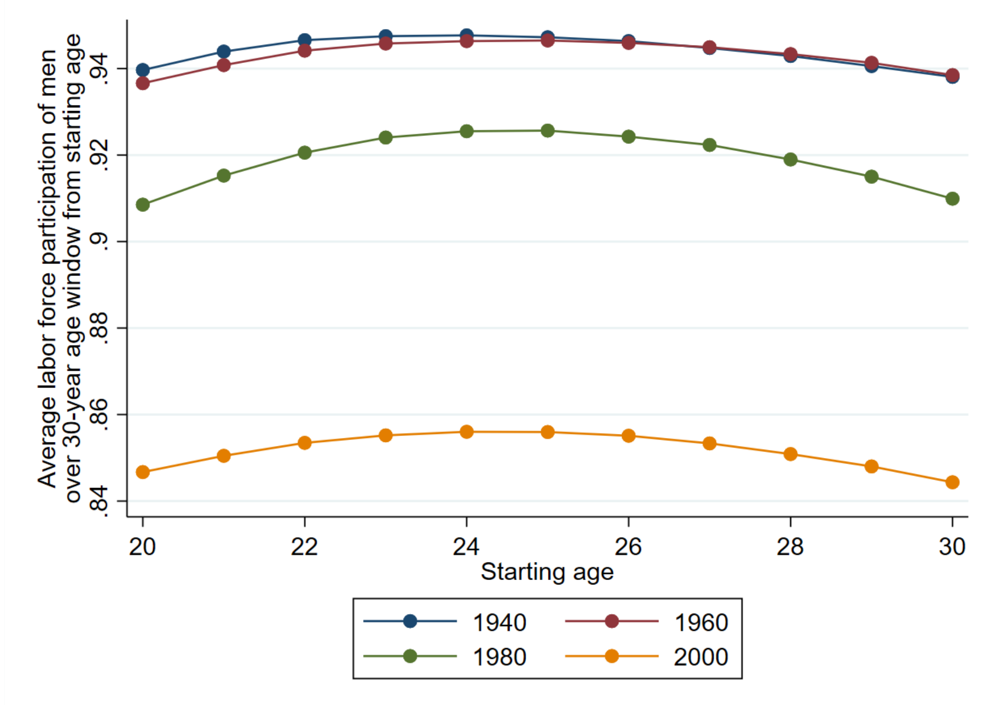
Notes: This figure plots the weighted average of labor force participation among men aged between starting age and (starting age+29). It shows that 25-54 is an appropriate age range for the economically active years of one's life and that this observation is quite stable over time.

Figure A4: Histogram of empirical market abilities



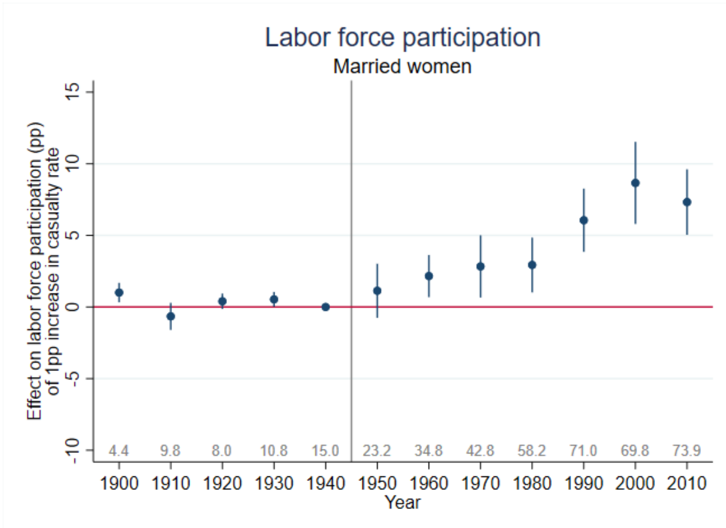
Notes: This figure plots the histogram of the market abilities of all working individuals with wage data, where the market abilities are inferred using the model structure as outlined in step 2 of section 4.3.

Figure A5: Average labor force participation of men over 30-year window from starting age

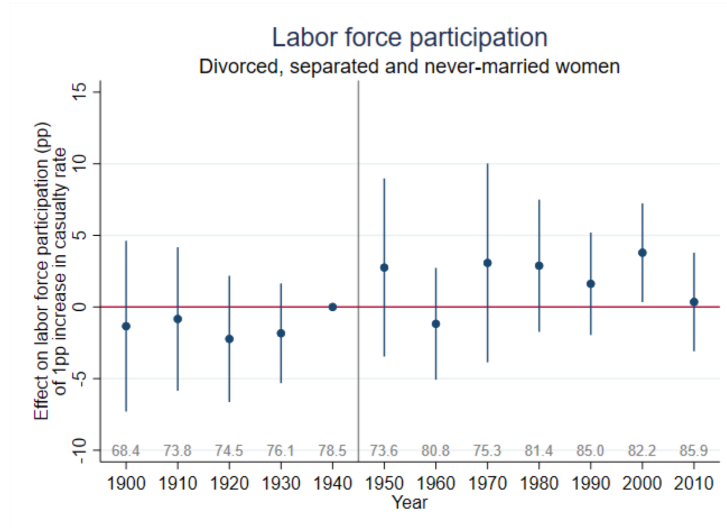


Notes: This figure plots the weighted average of labor force participation among men aged between starting age and (starting age+29). It shows that 25-54 is an appropriate age range for the economically active years of one's life and that this observation is quite stable over time.

Figure A6: The effect of WW2 draftee casualty rates on various outcomes

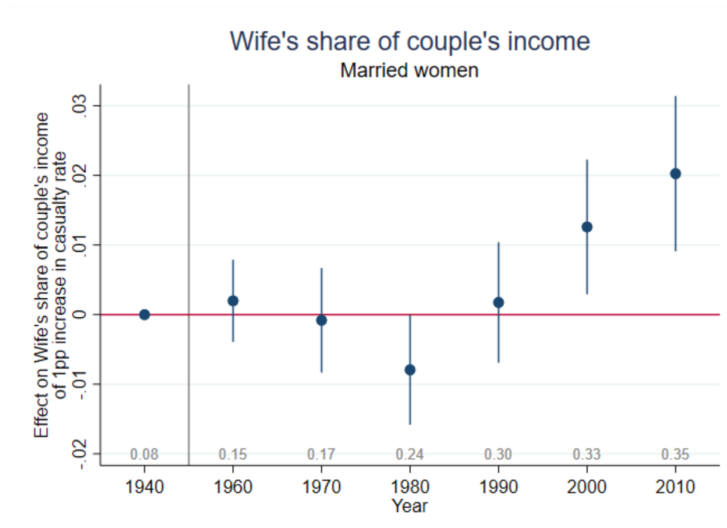


(A)

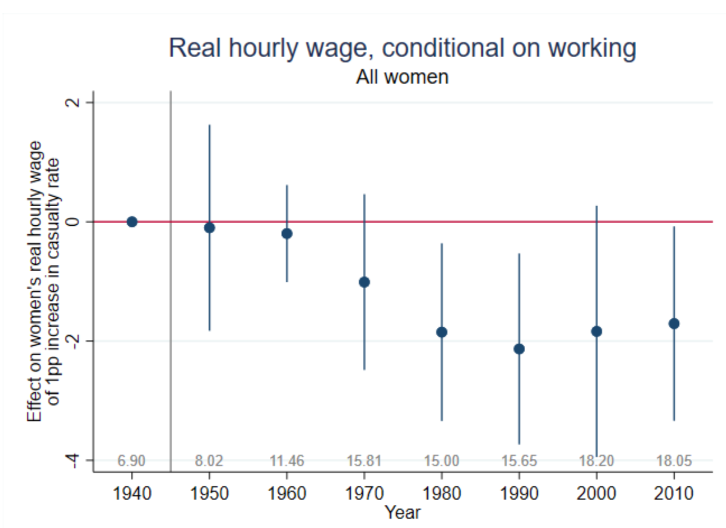


(B)

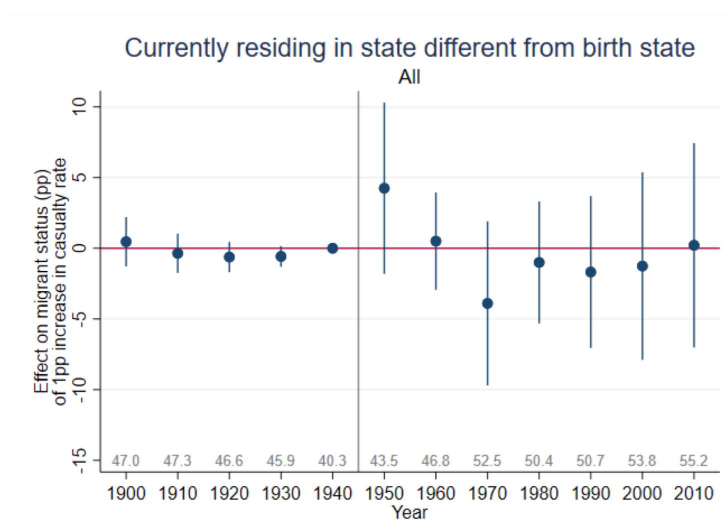
(C)



(D)

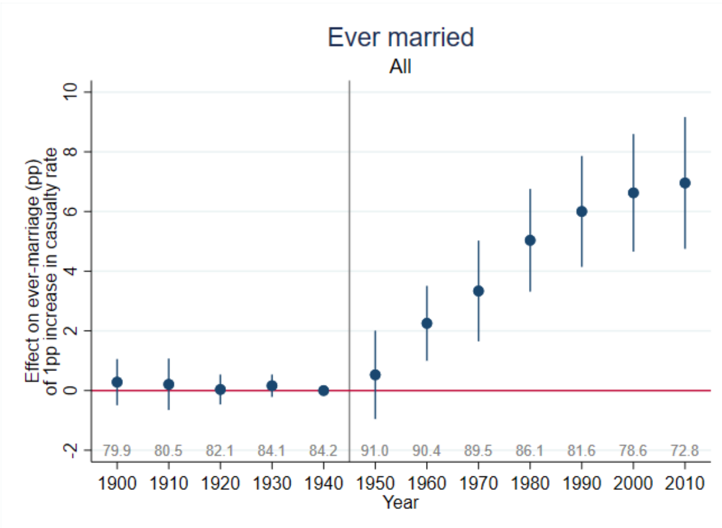


(E)

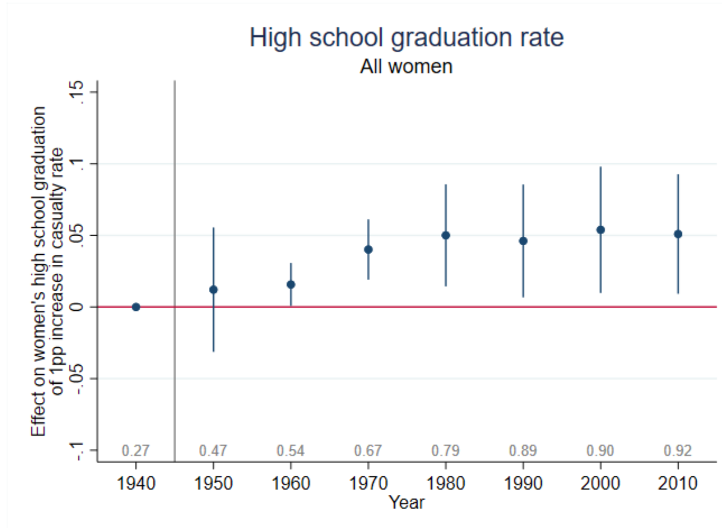


(F)

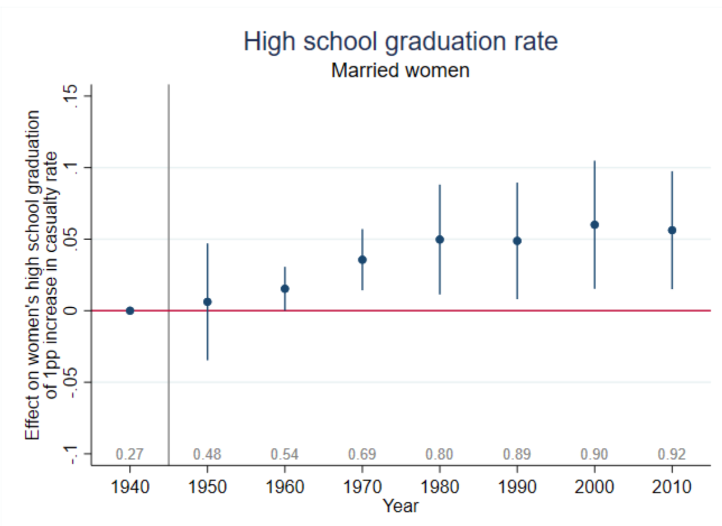
Figure A6 (continued): The effect of WW2 draftee casualty rates on various outcomes



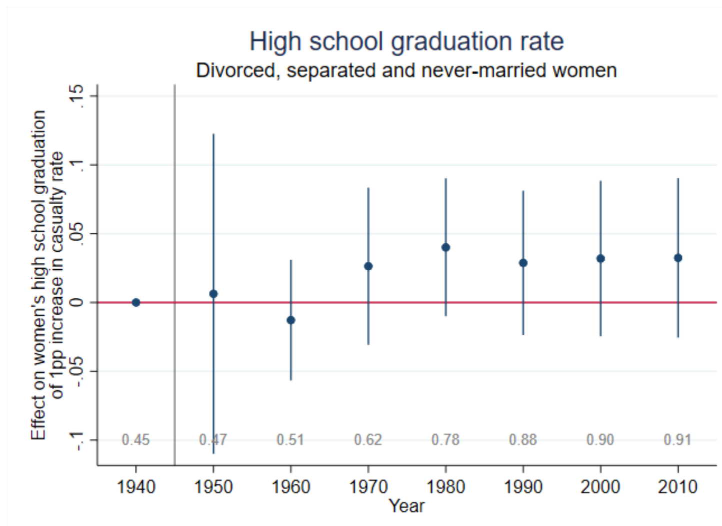
(G)



(H)



(I)



(J)

Notes: This figure plots the difference-in-differences coefficients from estimating equation (21) for various outcome variables.

Appendix B 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		

Notes: 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 if she has a husband capable of supporting her, in the Gallup Polls and the General Social Survey. The specific attitudinal survey question of interest 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 belonging to each variable, 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,∞)	[13,15]	[16,∞)	16	[17,∞)	[17,∞)	[17,∞)
6		[16,∞)		[17,∞)			

Notes: 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

Notes: 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

Notes: 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 detailed in Table A2 of the Appendix, schooling level is a categorical variable with higher values representing greater numbers of 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	-1.59	-0.64	0.00	0.65	1.58	
	Female	-1.70	-0.48	-0.08	0.46	1.81	
1960	Male	-1.07	-0.34	-0.49	-0.48	0.95	1.43
	Female	-1.27	-0.43	-0.75	-0.70	0.95	2.19
1970	Male	-0.88	-0.48	-0.71	0.66	1.41	
	Female	-1.09	-0.96	-0.95	0.83	2.16	
1980	Male	-0.75	0.07	-1.07	-0.33	0.90	1.18
	Female	-0.98	-0.51	-1.30	-0.14	1.19	1.73
1990	Male	-1.09	-1.24	-0.64	0.77	2.20	
	Female	-1.44	-1.33	-0.69	1.00	2.44	
2000	Male	-1.32	-1.63	-0.50	0.91	2.55	
	Female	-1.37	-1.59	-0.53	1.03	2.46	
2010	Male	-1.54	-1.62	-0.64	1.02	2.79	
	Female	-1.58	-1.45	-0.72	1.15	2.61	

Notes: This table reports the estimates of C^s_g, the direct cost minus the expected utility from intra-household resource allocation, by sex and decade. These estimates are not comparable by decade because the schooling level categories differ by decade, except for 1990, 2000, and 2010.

Table A6: The effect of WW2 draftee casualties on female labor force participation (percentage points)

	DID		DID with 1940 county controls		Synthetic DID
	continuous	binary	continuous	binary	binary
	(1)	(2)	(3)	(4)	(5)
1950	2.54***	2.13**	1.01	0.42	3.43***
	(0.98)	(0.86)	(0.83)	(0.87)	(0.15)
1960	2.13***	1.64**	0.66	0.53	4.07***
	(0.75)	(0.67)	(0.75)	(0.69)	(0.13)
1970	2.69**	2.03*	-0.09	-0.07	5.29***
	(1.11)	(1.09)	(-1.38)	(-1.25)	(0.47)
1980	4.03***	3.04***	1.64	0.55	6.81***
	(1.03)	(0.83)	(1.08)	(0.80)	(0.29)
1990	7.11***	5.57***	3.40***	1.76**	7.98***
	(1.15)	(0.94)	(1.16)	(0.87)	(0.33)
2000	9.78***	7.77***	5.08***	2.98***	9.50***
	(1.46)	(1.20)	(1.41)	(1.12)	(0.37)
2010	7.50***	6.23***	3.52***	2.41***	8.41***
	(1.04)	(0.92)	(1.01)	(0.93)	(0.36)
N	2,096,633	2,096,633	2,028,727	2,028,727	2,096,633

Notes: This table reports the decade-specific difference-in-differences coefficients of female labor force participation on WW2 draftee casualties by county. Draftee casualties are considered either as the raw continuous variable or as a binary variable equal to 1 if casualty rates are above the median. When casualties are continuous, the coefficients amount to percentage point changes in married women's labor force participation for every 1 percentage point increase in casualty rates. When casualties are binary, the coefficients amount to percentage point changes in married women's labor force participation for being in above-median counties relative to below-median counties. Columns (1) and (2) reports the result of estimating equation (21). Columns (3) and (4) add as controls 1940 county characteristics that predict casualty rates, interacted with decade dummies, in order to address the nonrandomness of casualty rates. Column (5) reports the coefficients from applying the synthetic difference-in-differences methodology (Arkhangelsky et al., 2019) to further allay concerns over *level* differences in the pre-WW2 period affecting the future trajectory of various outcome variables.

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

Notes: 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 C Model

C.1 General form of the utility function

The general form of the utility function is given by

$$u_i(Q, C_i, L_f, L_m) = H\Big(f(Q)C_i - r(Q)\big[\tau_{Fi}w_f L_f + \tau_{Mi}h_m(1-L_m)\big] + g_i(Q)\Big)$$

where the following conditions hold:
<u>**Conditions**</u>

C1) H is strictly increasing and strictly concave $\Rightarrow$ Individuals are risk-averse

C2) $(H')^{-1}$ is homogeneous or logarithmically homogeneous[38,39,40] $\Rightarrow$ Efficient risk-sharing implies spouses split the sum of individual "inverse-H" utilities[41] in an affine way.

C3) $2p(f')^2 - p\cdot f\cdot f'' + \big[\tau_F w_f L_f + \tau_M h_m(1-L_m)\big](r''f' - r'f'') - f'g'' + g'f'' > 0$, where $\tau_F \equiv \tau_{Fm} + \tau_{Ff}$, $\tau_M \equiv \tau_{Mm} + \tau_{Mf}$, and $g(Q) \equiv g_m(Q) + g_f(Q)$ $\Rightarrow$ Both spouses' indirect utility functions are increasing in the consumption of the public good *and* the optimal public good consumption is increasing in $M - \tilde{A}$.

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 nontraditional working arrangements, 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 problem is equivalent

[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).

[40] Mozzocco (2004) 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:

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

[41] By "inverse-H" utilities, I mean $H^{-1}(u_i)$, or the part of the utility function inside $H(.)$.

to solving

$$\max_{Q,C,Y,B} (Q+Y)\big(C+B-\tau_F w_f L_f - \tau_M h_m(1-L_m)\big)$$

s.t.

$$pQ + C = w_m L_m + w_f L_f$$

$$pY + B = h_m(1-L_m) + h_f(1-L_f)$$

where Y is the non-rival, public component of home production (e.g. cleaning of communal area, or food preparation for children) and $B \equiv B_m + B_f$ 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: On the perfect substitutability of home-produced goods and market goods

It is difficult to believe that home-produced goods and market goods are perfectly substitutable. 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 can be formulated as

$$u_i(Q, C_i, L_f, L_m) = H\Big(f(Q)C_i - r(Q)\big[a_i w_i L_i + \tau_{Fi} w_f L_f + \tau_{Mi} h_m(1-L_m)\big] + g_i(Q)\Big)$$

where a_i denotes i's preference for leisure/home consumption time, measured proportionally to i's market wage.[42] Ultimately, the optimal labor supply decisions would come down to

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

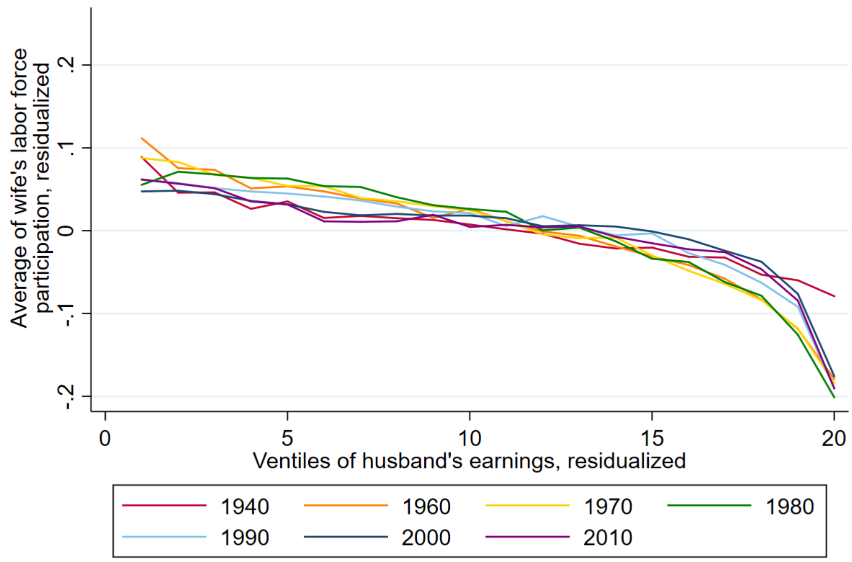
C.4 Model discussion: On the independence of husbands' and wives' labor supply decisions

In the real world, there is dependence in husbands' and wives' labor supply decisions (L_m and L_f) for many reasons, including specialization and diminishing marginal utility of household market consumption. The TU nature of the utility function in this model, along with the perfect substitutability between market goods and home produced goods, however, precludes such dependence.

[42] 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.

However, even in this model there is some degree of dependence, to the extent that wages and home productivities are modeled to be couple-specific; they depend on the schooling levels of both spouses, as well as their shared family composition characteristics (a vector of family size, number of children under 18, and number of children under 5).[43] For example, let us say a man with a master's degree and a woman who dropped out of high school are married. Empirically, the wife in this household is highly unlikely to work in the labor market, whereas a female high school dropout married to another high school dropout is more likely to work in the labor market. Thus there must be interdependence between husbands' and wives' labor supply decisions. But I can model their labor decisions as independent choices while having *market wages and home productivities to be match-specific*.

Figure A7: Average of wives' labor force participation by ventiles of husbands' earnings



Notes: This figure plots data derived from the sample of married women whose husbands are earning strictly positive wages. The horizontal axis denotes the first to the twentieth ventile of residuals from the regression of the husbands' earnings on age, race, years of schooling, spousal years of schooling, number of children, number of children under 5, family size dummies and county dummies. The reason for adding these particular controls is because these individual characteristics together define a group in the model. On the vertical axis is the weighted average of the residuals from the regression of the wives' labor force participation indicators on the same variables. All regressions are run separately for each decade.

Moreover, the concern of ignoring the dependence between L_m and L_f when evaluating the effects of changing gender norms over time is alleviated if the reasons behind the dependence are fixed over time. Therefore, I plot in Figure A7 the average of wives' labor force participation rates by ventiles of husbands' incomes, decade by decade, since husbands' earnings can be thought to affect households'

[43] Details about wages and home productivities are in subsection 3.1.

incentives for division of labor and marginal utilities of market consumption.[44] The pattern is very similar over time, except for the first and last ventile. This result demonstrates that drivers of the dependence of L_m and L_f might be more or less similar over time.

C.5 Derivation of marriage market equilibrium

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

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

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,

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

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,

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

[44] Both husbands' earnings and wives' labor force participation are residualized for race, schoolings of both spouses, number of children, number of children under 5, and family size.

Appendix D Parameter Inference

D.1 Deriving 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 $\times$ sex $\times$ education pair $\times$ 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.[45]

[45] 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.