The Role of Social Norms in Gender Division of Childcare and Fertility: A Quantitative Analysis of US-Japan Comparison

Tomoaki Kotera[♭], Yuta Okamoto[♯], and Daishin Yasui[♯]

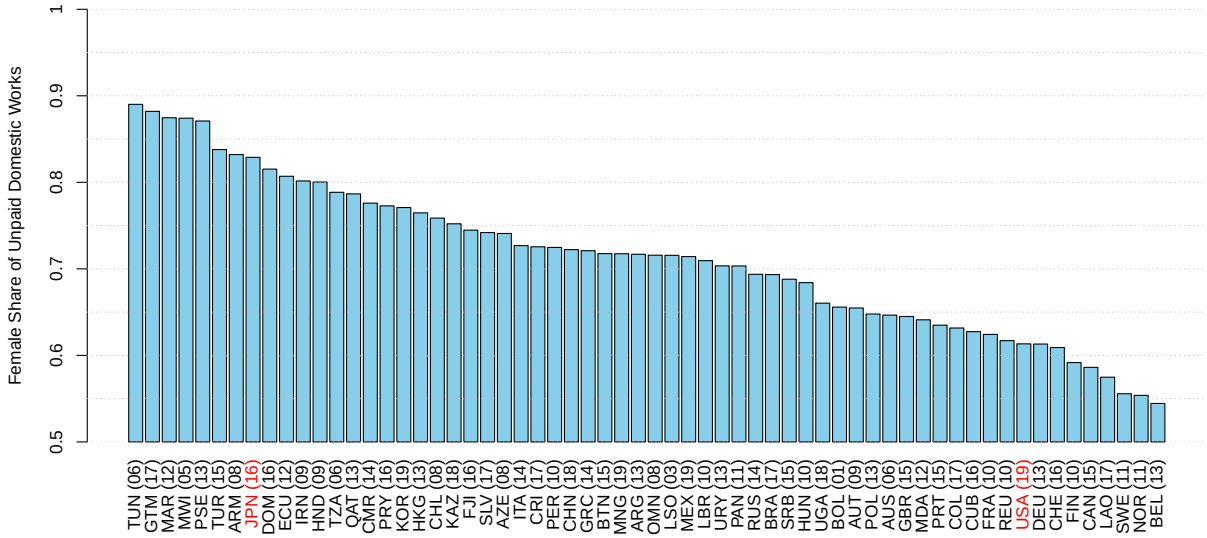
[♭]Department of Economics, Otaru University of Commerce
[♯]Graduate School of Economics, Kyoto University

September 14, 2025

Abstract

The division of childcare between the husband and wife substantially varies by country. In Japan, the burden is heavily biased toward women; this is not the case in the US. US and Japanese microdata reveal that only about 30% of the difference in marital division between the two countries can be explained by observable factors that are considered important determinants of marital division, such as differences in wages, education level, and employment status between married couples. We develop a model explaining these differences in marital division between the two countries based on the differences in social norms on the gender role in childcare, and calibrate it to the US and Japan. Our counterfactual analyses show that the presence of such social norms creates significant differences in fertility rates between the US and Japan, especially among couples where wives' wages are relatively high. This study provides a novel insight on which types of households are more likely to have more children by considering couples classified according to the wage ratio of the wife and husband.

Figure 1: Share of Women's Burden in Unpaid Domestic Work



Note: The average hours spent on unpaid domestic work by women, divided by the sum of the average hours spent by men and women, are shown. The survey years are in parentheses. The survey samples consist of individuals aged 15 and older. Source: *Minimum Set of Gender Indicators* by the United Nations.

Introduction

The sharing of childcare and housework between married couples substantially differs by country.[1] Figure 1 exhibits, for 60 countries, the ratio of the women's time to the total unpaid domestic work time of the couple in each country.[2] It is difficult to explain this large variation in the gender division of labor across countries solely in terms of cross-country differences in the gender gap in the labor market. According to the standard theory of family economics, the wage gap between married couples plays a central role in determining their division of childcare: the higher the husband's wage relative to the wife's, the higher the wife's share of the childcare burden, and vice versa. Figure 2a sorts households by the wage ratios of wives and husbands for Japan, where wives bear a

[1] Strictly speaking, childcare and housework should be distinguished. However, as will become apparent later, the gender shares of childcare and housework do not exhibit a large difference between in the data on married couples with children in the US and Japan used in this study. Furthermore, some housework, such as meal preparation and laundry, are activities that are difficult to completely separate from childcare. Thus, while our primary interest is childcare, we focus on both activities classified as childcare and housework in the data. Therefore, the terms "childcare" and "housework" are used without any distinction if there is no risk of confusion.

[2] Although this paper is concerned with the division of childcare and housework among married couples, due to limitations on data availability, Figure 1 compares the average burden of all unpaid domestic work between men and women aged 15 and older, and is not just limited to married couples. The main analysis is conducted using microdata from the US and Japan, where the analysis focuses on the division of childcare and housework among married couples.

relatively large share of the childcare burden, and the US, where they do not. It plots the average share of childcare burden borne by wives for each wage ratio.[3] As theoretically predicted, as the wife's wage ratio increases, the wife's share of the childcare burden declines in both countries. Figure 2b plots the distribution of the ratio of wages of wives to their husbands for the US and Japan. More households in the US have wives with relatively higher wages than in Japan. This indicates that US-married women are less disadvantaged in the labor market compared to Japanese married women. However, such a difference in household composition alone cannot explain the average difference between the two countries in Figure 1. Even when households with the same wife wage ratio are compared, the ratio of childcare burden for wives in Japan is considerably higher than that in the US. This paper develops a model which can explain this difference in the gender division of childcare between the US and Japan by the differences in social norms. Based on the model, we quantify the impact of the presence of social norms on differences in the overall level and distribution of fertility among households between the two countries.

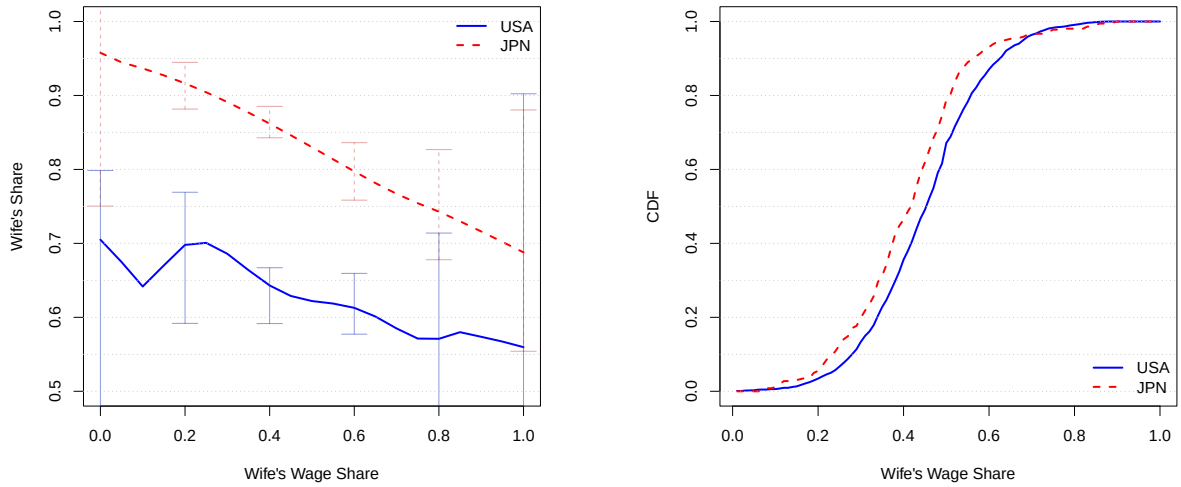
The East Asian region, including Japan, is known for its traditional view of the gendered division of labor, with men working outside the home and women taking care of the home. Figure 1 lists 60 countries and regions by the percentage of women in childcare and housework, with Japan in 8th place, South Korea in 17th place, Hong Kong in 18th place, and China in 27th place. In Japan, there is a term called "Ryōsai Kenbo." This term, which means "Good Wife, Wise Mother," began appearing in the mass media in the late 19th century and was used until after World War II to describe the ideal image of Japanese women. Specifically, women in Japan were expected to master domestic skills, such as sewing and cooking, and acquire the moral and intellectual ability to raise strong and intelligent children for the nation. Furthermore, until the 1990s, Japan adopted an educational curriculum that encouraged a gender division of labor in schools. In junior high and high school, only women were required to take home economics, where they learned skills related to cooking, sewing, managing the household budget, and interacting with children and the elderly. Although the Ryōsai Kenbo ideology rapidly lost support in the latter half of the 20th century and the aforementioned educational curriculum was abolished in the 1990s, Figure 2a suggests that this ideology and education still underlie the thinking of many Japanese people, and are the aspects on which the gender division of labor is actually based.[4]

In our model, each married couple cooperatively determines intra-household allocation, including the division of childcare time. On the one hand, the relationship between

[3] In calculating women's share of the burden in Figure 2a , activities classified as child care and housework are included. However, what we really intend to focus on are activities that increase the time spent as the number of children increases. Limiting the calculation to only those activities classified as childcare is too narrow. Meanwhile, including activities classified as housework is too broad. As ths Online Appendix Section S1.1 shows, our results do not substantially differ whether we limit ourselves to activities classified as childcare or include those classified as housework.

[4] Hara and Rodriguez-Planas (2025) find that between cohorts educated before and after the Japan's 1990s education curriculum reform, gender gap outcomes, as measured by gender differences in time spent in home production, time spent in work-related activities, employment status, and earnings, differ. This suggests the existence of norms and beliefs about the division of gender roles behind the actual observed gender division of labor.

Figure 2: Female's Share of Domestic Work & Relative Wage



(a) Share of Domestic Work by Relative Wage

(b) Distribution of Relative Wage

Note: We focus on couples with cohabiting children where both the husband and wife are working. (a) The wife's share in the U.S. is estimated as $\hat{\mathbb{E}}[Y^f|R]/(\hat{\mathbb{E}}[Y^f|R]+\hat{\mathbb{E}}[Y^m|R])$, where Y^f and Y^m denote the time spent on domestic work by wives and husbands, respectively, and R is the wife's wage share. In Japan, the share is estimated as $\hat{\mathbb{E}}[Y^f/(Y^f+Y^m)|R]$. All regression functions are estimated using local linear regression with IMSE-optimal bandwidths and a Gaussian kernel. The 95% (pointwise) confidence intervals (CIs) are constructed using the robust bias-correction method of Calonico et al. (2018). (b) The empirical distribution functions are shown. Source: *American Time Use Survey* (2012, 2019) and *Japan Household Panel Survey* (2012, 2019).

childcare production technology and the spousal wage gap determines the efficient division of childcare time for that couple. On the other hand, the desirable division of childcare time is dictated by social norms regarding gender roles, with extant incentives to get closer to this desired division. Thus, each couple faces a tradeoff between efficiency and social norms in determining its division of childcare time. The greater the distance between the efficient division of labor determined by the spousal wage gap and the socially desirable division of labor, the greater the psychological cost incurred by the couple due to the existence of the social norm. Thus, differences in social norms cause differences in the overall fertility rate. More importantly, they cause differences in the distribution of the number of children; that is, which wage combinations couples have more and fewer children.

Our study is related to work on the role of social norms and culture in labor-supply and fertility decisions, including Fernández et al. (2004), Fernández and Fogli (2006, 2009), Bertrand et al. (2015), Kuehnle et al. (2021), de Silva and Tenreyro (2020), Bertrand et al. (2021), Myong et al. (2021), and Beach and Hanlon (2022).[5] This paper is most closely re-

[5]Similar approaches to studies that assume social norms also exist, which assume preference externalities. Such studies, including de Laat and Sevilla-Sanz (2006) and Kim et al. (2024), typically consider

lated to Myong et al. (2021) in that we share the motivation to analyze the impact of social norms regarding gender roles in childcare on fertility behavior. They compare the US and South Korea, and their analysis assumes that the US is an economy where there are no social norms and couples optimally determine the division of childcare, while South Korea is an economy where all households are faced with a fixed gender share of childcare burden according to social norms. In contrast, we compare the US and Japan, assuming that the US is an economy with social norms in which men are relatively cooperative in child-rearing, while Japan is an economy with social norms in which men are uncooperative in child-rearing. Our approach has the advantage of explaining the following. First, even in the US, wives are the primary childcare providers, even if their wages are considerably higher than those of their husbands. Thus, the US also cannot be considered free from social norms. Second, although women bear a much greater burden in Japan than in the US, the elasticity of the ratio of the childcare burden to the ratio of wives' to husbands' wages is not lower than that in the US: it is not that one is rigid and the other elastic, but that both are elastic to some extent but at different levels.

One contribution of our work is that it explicitly takes into account the effect of gender sharing of childcare burden in determining the number of children. Most theoretical literature on fertility choice has relied on the assumption that only women are responsible for childcare and exclusively focused on the effect of female wages on the number of children.[6] This simplistic view can be justified when analyzing situations in which women are responsible for almost all childcare. However, a notable change in the labor market observed in many countries since the late 20th century has been the increase in female participation in the workforce and male participation in the home. Thus, focusing only on female childcare may be misleading when considering childcare issues in such changed societies. Doepke and Kindermann (2019) and Myong et al. (2021) are among the few studies that explicitly consider the effect of gender childcare sharing on the number of children. Doepke and Kindermann (2019) assume that the gender share of childcare burden is given exogenously and that this share varies across countries. Myong et al. (2021) compare an economy with an efficient division of labor for childcare based on wage ratios between husbands and wives to a rigid one where the burden share is given exogenously. In contrast, in this study, couples make marital childcare sharing decisions while facing a tradeoff between efficient sharing and social norms. Our model has two main advantages. First, it can simultaneously account for not only regularities between different economies but also those among different households within the same economy. Second, it allows us to analyze economies with different social norms within a unified framework.[7]

Another contribution of this study is to shed light on the role of the wage ratio between wives and husbands in determining their share of childcare and the number of children. While the literature on intra-household time allocation, such as female labor

situations in which a society's average behavior affects each individual's behavior.

[6]For instance, Jones et al. (2010) and Doepke et al. (2023) provide comprehensive reviews of the theoretical literature on fertility choice.

[7]Our model abstracts from important features in the model of Doepke and Kindermann (2019), such as bargaining between spouses and lack of commitment, and those in the model of Myong et al. (2021), such as marriage decisions.

supply and home production, has focused on the wage ratios of wives and husbands (e.g., Albanesi and Olivetti, 2009; Knowles, 2013), fertility choice literature focuses on, wives' wages, with little attention to the wage ratios of wives and husbands.[8] This paper analyzes the effect of the existence of social norms regarding gender childcare sharing on fertility choice, which is shown to have a large effect on the distribution of fertility (what type of household has how many children) and overall fertility rate. Furthermore, the distributional effect is small when households are classified by household income and large when classified by the wage gap between married couples. When considering the effects of policies such as pro-natal policies, it is important to consider not only their impact on overall fertility, but also what types of households are affected, and this study provides a new insight on such distributional effects.

1 Motivating Evidence

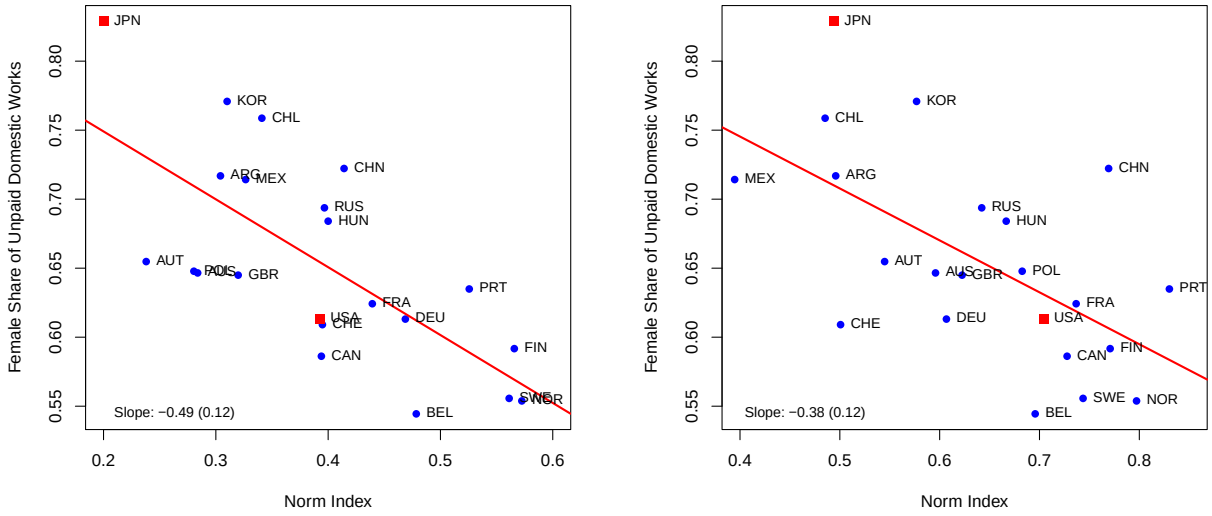
The main assumption of our analysis is that the social norms prevalent in an economy play an important role in how the childcare burden is typically shared between married couples in that economy. Here, we provide several suggestive evidences supporting this assumption.

Using cross-country aggregate data, we first assess the validity of focusing on social norms by examining the correlation between attitudes toward gender roles and actual gender division of labor in the home. We use data from the Minimum Set of Gender Indicators provided by the United Nations, and estimate the actual gender share of domestic work as described in Figure 1. For attitudes toward gender roles, we construct a measure that captures such attitudes based on the International Social Survey Programme (ISSP) 2012 (ISSP Research Group, 2016). The survey respondents are asked whether women with children below school age or who have started school, should work full-time, part-time, or stay home. We assign a value of 1, 0.5, and 0 to respondents who said they should work full-time, part-time, and stay at home, respectively, and calculate the average for each country. Thus, our constructed measure of attitudes toward gender roles takes values between 0 and 1, with smaller values indicating stronger social norms that women with children should be responsible for domestic work.

To examine how the division of childcare among married couples is related to their attitudes toward gender roles, in Figure 3, we plot the percentage of childcare sharing actually chosen by married couples against the measure of attitudes toward gender roles constructed by the above method. Figures 3a and 3b use the attitude measure for mothers with preschool-aged children and children who have started school, respectively. We observe a clear negative correlation regardless of which attitude measure is used, suggesting that in countries with smaller social norms that women with children should stay home and not work outside the home, women are more likely to have a smaller share of

[8] In the models of Galor and Weil (1996) and Kimura and Yasui (2010), the wage ratio between wives and husbands plays a central role in determining fertility. However, in their studies, the division of childcare between husbands and wives is not considered; basically, only wives are responsible for childcare.

Figure 3: Social Attitudes and Domestic Work Division



(a) The Youngest Kids under School Age

(b) After the Youngest Kids Starts School

Note: Wives' share of domestic work is calculated as in Figure 1. For the norm index, we assign a value of 1 for full-time, 0.5 for part-time, and 0 for stay-home, and compute its average for each country. The linear fits are shown in red. Source: ISSP (2012) and *Minimum Set of Gender Indicators* by the United Nations.

the housework burden. This pattern is robust to the choice of different specifications for the attitude measure (see Online Appendix Section S1.2 for details).

Next, we take Japan and US as the countries where childcare is strongly biased toward women and relatively unbiased, respectively, and examine whether the difference observed in Figure 2a is simply brought about by observable labor market factors, such as differences in the education levels and employment status between Japanese and American couples, or whether there is room to attribute it to social norms. For this, we use two microdata sources: the 2012 and 2019 waves of the Japan Household Panel Survey (JHPS) and the American Time Use Survey (ATUS). For the JHPS, we use only the new refreshment sample data added in 2012 and 2019, and exclude individuals surveyed in earlier waves in order to minimize potential attrition bias. In our analysis, we focus on opposite-sex couples in which both partners are between the ages of 25 and 55, are employed, have children living with them, and do not live with their parents. We also exclude samples that gave birth within one year of the survey year because the recorded content may not accurately reflect the information we would like to obtain (e.g., the recorded time use may reflect the period after childbirth, whereas the reported earnings and working hours may correspond to the period before childbirth).

Using the ATUS data, we first train the following predictive regression model:

$$Y_i^j = \beta_0^j + \beta_1^j R_i + Z_i' \gamma^j + \varepsilon_i,$$

Table 1: Difference in Share of Time Spent on Domestic Work between the US and Japan

	Childcare			Total Domestic Work		
	JPN	JPN (CF)	USA	JPN	JPN (CF)	USA
Share	0.834	0.707	0.647	0.861	0.690	0.634

Note: Columns JPN and USA show the actual shares, while the estimated counterfactual shares are shown in the column JPN (CF). Source: ATUS (2012, 2019) and JHPS (2012, 2019).

where $j \in \{m, f\}$ with m representing male and f representing female, Y_i^j denotes the daily time spent on childcare or housework (in minutes) by gender j in couple i, R_i is the wife's wage share, and Z_i is a vector of other predicting variables. The vector Z_i includes the number of children living together, the age of the youngest child, the indicator if the family income is below the median family income for the country, the ages of the couple and their interaction, couple's education years and their interaction, and the indicators of full-time worker and their interaction. Since only one person per household is interviewed in the ATUS data, the regressions are run separately for husbands and wives.

After training the above model for the US, we use the JHPS data to calculate the hours of childcare and housework in the counterfactual scenario that Japanese couples would choose if they were in the predicted US economy:

$$\hat{Y}_i^j = \hat{\beta}_0^j + \hat{\beta}_1^j R_i + Z_i' \hat{\gamma}^j.$$

The JHPS data includes information about both husband and wife in each couple, allowing us to compute the average actual and counterfactual share by the sample mean of $Y_i^f/(Y_i^f + Y_i^m)$ and $\hat{Y}_i^f/(\hat{Y}_i^f + \hat{Y}_i^m)$, respectively. The US actual share is computed as $\hat{\mathbb{E}}[Y_i^f]/(\hat{\mathbb{E}}[Y_i^f] + \hat{\mathbb{E}}[Y_i^m])$. Table 1 provides the actual share in Japan, counterfactual share for Japanese couples, and actual share in the US. First, the results are almost the same whether we calculate the time spent only on childcare, or childcare and housework combined. Second, the wife's share of childcare and housework in counterfactual Japanese couples is considerably lower than the actual Japanese wife's share. This suggests large differences which cannot be explained by differences in (R_i, Z_i'). Specifically, even if we control for differences in (R_i, Z_i'), 66-74% of the difference in the gender burden at home between the two countries remains unexplained.

Together, the results in Figure 3 and Table 1 suggest that the differences in the gender division of labor between the two countries may reflect differences in the social norms prevalent in the two countries. Therefore, next, we construct a model in which the unexplained part in Table 1 is brought about by differences in social norms, and analyze how the existence of different social norms affects the differences in fertility behavior in the two countries.

2 Main Mechanism

In this section, we build a simple model to highlight the main mechanism of this paper. We then test whether some predictions derived from that simple model are consistent

with the data we use.

2.1 Simple Model

We start with a simple model to obtain the sharpest predictions. The next section deals with extensions aimed at building a quantitative model that can be matched to the data.

A couple comprises a woman f and man m with different wages, who cooperate in making decisions regarding birth, intra-household time allocation, and intra-household consumption allocation. When making these decisions, the couple are influenced by a social norm regarding the ratio of wife and husband childcare sharing: there is a socially appropriate childcare sharing ratio that reflects the roles which society expects each gender to play, and each household tries to avoid moving away from that ratio.

Raising a child requires a time investment $\phi \in (0,1)$, which is divided between the couple: $\phi = x_f + x_m$, where x_j is childcare time of gender $j \in \{f,m\}$. We assume that wife childcare and husband childcare are perfect substitutes, although this assumption is relaxed in the quantitative analysis, and efficiency between the two does not differ. Denote the wife's share of childcare time by $r \equiv x_f / \left(x_f + x_m\right)$, and thus, $x_f = r\phi$ and $x_m = (1-r)\phi$.

The utility of gender j is given by the following:

$$u_j\left(c_j, l_j, n, r\right) = \ln c_j + \varsigma \ln l_j + \gamma n - n \cdot \frac{\psi\left(r - \tilde{r}\right)^2}{2},$$

where c_j is consumption of gender j, l_j is leisure of gender j, and n is fertility. To simplify the analysis, we consider a discrete fertility choice where couples decide a single child or not, $n \in \{0,1\}$.[9] The social norm is represented as a ratio $\tilde{r}$ of wife's childcare burden that reflects society's expectations of women's role in childcare. The couple suffers a utility loss when deviating the ratio. Parameters ς, γ, and ψ represent the degree of importance attached to children, leisure, and social norms, respectively. The couple's budget constraint is the following:

$$c_f + c_m = w_f\left(1 - l_f - r\phi\right) + w_m\left(1 - l_m - (1-r)\phi\right), \tag{2.1}$$

where w_f and w_m are wife's and husband's wages per unit of time, respectively. We assume cooperative bargaining in which the couple solves a Pareto problem with welfare weights θ_f and θ_m, where $\theta_f + \theta_m = 1$, for the wife and the husband, respectively. Thus, the objective function of the couple is given by $\theta_f u_f\left(c_f, l_f, n, r\right) + \theta_m u_m\left(c_m, l_m, n, r\right)$.

If the couple decides to have a child, $n = 1$, the couple maximizes

$$\theta_f\left(\ln c_f + \varsigma \ln l_f\right) + \theta_m\left(\ln c_m + \varsigma \ln l_m\right) + \gamma - \frac{\psi\left(r - \tilde{r}\right)^2}{2}$$

subject to (2.1).

[9] The utility derived from children is linear here. However, this is without loss of generality given that the number of children can only be 0 or 1. In the quantitative model, where the number of children can be chosen from any real number, we assume that marginal utility from children diminishes.

The wife's share of childcare time chosen by the couple in the interior case, r^I, is given by:

$$-\psi\left(r^I-\tilde{r}\right)-\frac{(1+\varsigma)\,\phi\,(2\omega-1)}{\omega\,(1-r^I\phi)+(1-\omega)\,[1-(1-r^I)\,\phi]}=0, \tag{2.2}$$

where $\omega \equiv w_f/\left(w_f+w_m\right)$ is the wife's relative wage to her husband (see Appendix A for the derivation of this condition). The first term captures the incentive to make their share closer to the socially desirable one: If $r^I<\tilde{r}$, the effect is to increase r^I; if $r^I>\tilde{r}$, the effect is to decrease r^I. The second term captures the incentive to allocate time efficiently in terms of maximizing income: If $\omega<1/2$, the effect is to increase r^I; if $\omega>1/2$, the effect is to decrease r^I. We can easily find that $r^I>\tilde{r}$ if $\omega<1/2$, $r^I=\tilde{r}$ if $\omega=1/2$, and $r^I<\tilde{r}$ if $\omega>1/2$.

Considering the case of the corner solutions, the couples' choice can be summarized as follows:

$$r^*=\begin{cases}1 & \text{if } \omega<\underline{\omega},\\ r^I & \text{if } \underline{\omega}\le\omega<\overline{\omega},\\ 0 & \text{if } \omega\ge\overline{\omega},\end{cases}$$

where

$$\underline{\omega}=\frac{(1+\varsigma)\,\phi-\psi\,(1-\tilde{r})}{[2\,(1+\varsigma)-\psi\,(1-\tilde{r})]\,\phi} \text{ and } \overline{\omega}=\frac{(1+\varsigma)\,\phi+\psi\tilde{r}\,(1-\phi)}{[2\,(1+\varsigma)-\psi\tilde{r}]\,\phi}.$$

If the wife's relative wage is sufficiently low, the wife takes care of the child alone. Conversely, if the wife's relative wage is sufficiently high, the husband takes care of the child alone. In the intermediate case, where the wage gap between husband and wife is not too large, they share the childcare.

Without social norm (i.e., in the case of $\psi=0$), the optimal division of childcare between husband and wife is

$$r^o=\begin{cases}1 & \text{if } \omega<1/2,\\ [0,1] & \text{if } \omega=1/2,\\ 0 & \text{if } \omega>1/2.\end{cases}$$

In the absence of the social norm, if the husband's wage is higher than the wife's wage, childcare is provided by the wife alone, and vice versa. This result comes from the perfect substitutablity between wife's and husband's childcare. The deviation of r^* from r^o is created by the existence of social norms.

We can obtain some comparative-static results regarding the sharing of childcare burden. The effect of an increase in the wife's relative wage on the wife's childcare burden is

$$\frac{dr^I}{d\omega}=-\frac{\frac{(1+\varsigma)\phi(2-\phi)}{\{\omega(1-r^I\phi)+(1-\omega)[1-(1-r^I)\phi]\}^2}}{\frac{(1+\varsigma)\phi^2(2\omega-1)^2}{\{\omega(1-r^I\phi)+(1-\omega)[1-(1-r^I)\phi]\}^2}+\psi}<0.$$

The higher the wife's relative wage, the smaller her childcare burden. The effect of an increase in the wife's burden of childcare considered socially desirable on the one actually

chosen is

$$\frac{dr^I}{d\tilde{r}} = \frac{\psi}{\frac{(1+\varsigma)\phi^2(2\omega-1)^2}{\{\omega(1-r^I\phi)+(1-\omega)[1-(1-r^I)\phi]\}^2}+\psi} \in (0,1).$$

Changes in social norms change the actual sharing ratio chosen in the same direction (i.e., $dr^I/d\tilde{r} > 0$), but the actual change is not as responsive as the change in social norms (i.e., $dr^I/d\tilde{r} < 1$). The effect of an increase in the degree of emphasis on social norms is

$$\frac{dr^I}{d\psi} = -\frac{r^I-\tilde{r}}{\frac{(1+\varsigma)\phi^2(2\omega-1)^2}{\{\omega(1-r^I\phi)+(1-\omega)[1-(1-r^I)\phi]\}^2}+\psi} \begin{cases} <0 & \text{if } \omega < 1/2, \\ =0 & \text{if } \omega = 1/2, \\ >0 & \text{if } \omega > 1/2. \end{cases}$$

The greater the emphasis on social norms, the greater the deviation from the efficient burden ratio in the absence of social norms: An increase in ψ has the effect of making r^I away from r^o. The effect of an increase in time spent on childcare is

$$\frac{dr^I}{d\phi} = -\frac{\frac{(1+\varsigma)(2\omega-1)}{\{\omega(1-r^I\phi)+(1-\omega)[1-(1-r^I)\phi]\}^2}}{\frac{(1+\varsigma)\phi^2(2\omega-1)^2}{\{\omega(1-r^I\phi)+(1-\omega)[1-(1-r^I)\phi]\}^2}+\psi} \begin{cases} >0 & \text{if } \omega < 1/2, \\ =0 & \text{if } \omega = 1/2, \\ <0 & \text{if } \omega > 1/2. \end{cases}$$

As the time required for childcare increases, income decreases and the marginal utility of consumption increases, leading to a focus on efficient burden sharing rather than social norms. Thus, an increase in ϕ has the effect of making r^I closer to r^o.

By comparing the value of having a child versus not having a child, we can consider which couples will have a child. The value of the couple in the case of $n = 1$ is

$$\begin{aligned} V(1) &= \theta_f u_f(c_f^*, l_f^*, 1, r^*) + \theta_m u_m(c_m^*, l_m^*, 1, r^*) \\ &= \theta_f(1+\varsigma)\ln\theta_f + \theta_m(1+\varsigma)\ln\theta_m + \varsigma\ln\varsigma - (1+\varsigma)\ln(1+\varsigma) - \theta_f\varsigma\ln w_f - \theta_m\varsigma\ln w_m \\ &\quad + (1+\varsigma)\ln\left\{w_f(1-r^*\phi) + w_m[1-(1-r^*)\phi]\right\} + \gamma - \frac{\psi(r^*-\tilde{r})^2}{2}. \end{aligned}$$

The value of the couple in the case of $n = 0$ is

$$\begin{aligned} V(0) &= \theta_f u_f(c_f^*, l_f^*, 0, r) + \theta_m u_m(c_m^*, l_m^*, 0, r) \\ &= \theta_f(1+\varsigma)\ln\theta_f + \theta_m(1+\varsigma)\ln\theta_m + \varsigma\ln\varsigma - (1+\varsigma)\ln(1+\varsigma) \\ &\quad -\theta_f\varsigma\ln w_f - \theta_m\varsigma\ln w_m + (1+\varsigma)\ln\left(w_f + w_m\right). \end{aligned}$$

The difference in value between having and not having a child is expressed as follows:

$$\begin{aligned} & V(1) - V(0) \\ = \; & (1+\varsigma)\ln\{\omega(1-r^*\phi) + (1-\omega)[1-(1-r^*)\phi]\} + \gamma - \frac{\psi(r^*-\tilde{r})^2}{2} \\ = \; & \begin{cases} (1+\varsigma)\ln(1-\omega\phi) + \gamma - \frac{\psi(1-\tilde{r})^2}{2} & \text{if } \omega < \underline{\omega}, \\ (1+\varsigma)\ln\{\omega\left(1-r^I\phi\right) + (1-\omega)\left[1-\left(1-r^I\right)\phi\right]\} + \gamma - \frac{\psi\left(r^I-\tilde{r}\right)^2}{2} & \text{if } \underline{\omega} \le \omega < \overline{\omega}, \\ (1+\varsigma)\ln[1-(1-\omega)\phi] + \gamma - \frac{\psi\tilde{r}^2}{2} & \text{if } \omega \ge \overline{\omega}. \end{cases} \end{aligned}$$

The couple has a child if and only if $V(1) - V(0) > 0$. By investigating $V(1) - V(0)$, we can understand what types of households are more likely to have a child and obtain the following three propositions as testable predictions.

Proposition 1. *For households in which wives bear a greater (smaller) burden of childcare, an increase in their relative wages decreases (increases) fertility.* ♣

Proof. The effect of changes in the wife's relative wage on fertility choices is captured by

$$\frac{\partial [V(1) - V(0)]}{\partial \omega} = \begin{cases} -\frac{(1+\varsigma)\phi}{1-\omega\phi} < 0 & \text{if } \omega < \underline{\omega}, \\ \frac{(1+\varsigma)\left(1-2r^I\right)\phi}{\omega(1-r^I\phi)+(1-\omega)[1-(1-r^I)\phi]} & \text{if } \underline{\omega} \leq \omega < \overline{\omega}, \\ \frac{(1+\varsigma)\phi}{1-(1-\omega)\phi} > 0 & \text{if } \omega \geq \overline{\omega}, \end{cases}$$

where we use the envelope theorem in the case of $\underline{\omega} \leq \omega < \overline{\omega}$. Since $dr^I/d\omega < 0$, we obtain

$$\frac{\partial [V(1) - V(0)]}{\partial \omega} \begin{cases} < 0 & \text{if } \omega < \hat{\omega}, \\ > 0 & \text{if } \omega > \hat{\omega}, \end{cases}$$

where $\hat{\omega}$ is the wife's relative wage at which her share of childcare burden is $1/2$. □

Proposition 2. *The prevalence of social norms that require women to bear a greater burden of childcare increases (decreases) fertility in households where husbands (wives) have higher wages.* ♣

Proof. The effect of changes in social norms on fertility choices is captured by

$$\frac{\partial [V(1) - V(0)]}{\partial \tilde{r}} = \begin{cases} \psi(1 - \tilde{r}) & \text{if } \omega < \underline{\omega}, \\ \psi\left(r^I - \tilde{r}\right) & \text{if } \underline{\omega} \leq \omega < \overline{\omega}, \\ -\psi\tilde{r} & \text{if } \omega \geq \overline{\omega}, \end{cases}$$

where we use the envelope theorem in the case of $\underline{\omega} \leq \omega < \overline{\omega}$. Since $r^I > \tilde{r}$ when $\omega < 1/2$, $r^I = \tilde{r}$ when $\omega = 1/2$, and $r^I < \tilde{r}$ when $\omega > 1/2$, we obtain

$$\frac{\partial [V(1) - V(0)]}{\partial \tilde{r}} \begin{cases} > 0 & \text{if } \omega < 1/2, \\ = 0 & \text{if } \omega = 1/2, \\ < 0 & \text{if } \omega > 1/2. \end{cases}$$

The social norm that women should be in charge of childcare positively (negatively) affects fertility choice in households where wives earn lower (higher) wages. □

Proposition 3. *A complementarity exists between women's wages and social norms that require male childcare participation in fertility decisions: the more pervasive the social norm that men should actively participate in child rearing is in an economy, the less likely that the fertility rate will decline when women's position in the labor market becomes stronger.* ♣

Proof. In the case of $\underline{\omega} \leq \omega < \overline{\omega}$, we have:

$$\frac{\partial^2 [V(1) - V(0)]}{\partial \omega \partial \tilde{r}} = \psi \frac{dr^I}{d\omega} = -\psi \frac{\frac{(1+\varsigma)\phi(2-\phi)}{\{\omega(1-r^I\phi)+(1-\omega)[1-(1-r^I)\phi]\}^2}}{\frac{(1+\varsigma)\phi^2(2\omega-1)^2}{\{\omega(1-r^I\phi)+(1-\omega)[1-(1-r^I)\phi]\}^2} + \psi} < 0.$$

Therefore, we obtain:

$$\frac{\partial^2 [V(1) - V(0)]}{\partial \omega \partial \tilde{r}} \begin{cases} = 0 & \text{if } \omega < \underline{\omega}, \\ < 0 & \text{if } \underline{\omega} \leq \omega < \overline{\omega}, \\ = 0 & \text{if } \omega \geq \overline{\omega}. \end{cases}$$

The negative sign of the cross derivative with respect to ω and $\tilde{r}$ immediately means the statement of this proposition. □

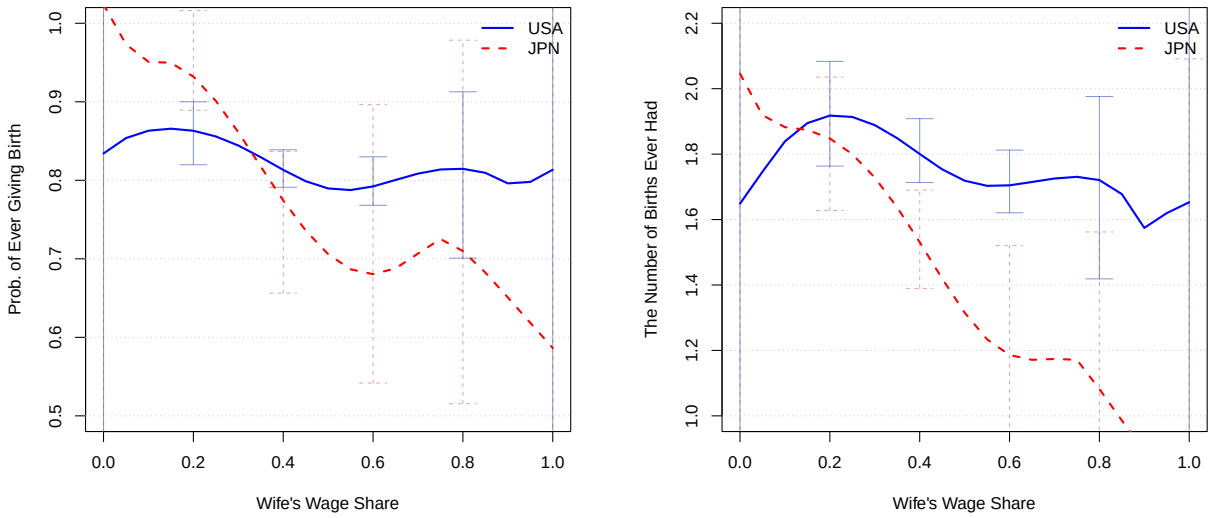
2.2 Testing Predictions

Here, we check the plausibility of the main mechanism of our model by comparing the predictions derived from Propositions 1-3 with US and Japanese data used in the quantitative analysis. To constrcut fertility variables, for Japan, we again use the JHPS 2012 and 2019 data; for the US, we use the fertility supplement of the Current Population Survey (CPS) 2012 and 2018, which are the waves closest to the JHPS 2012 and 2019. Based on these data, we nonparametrically estimate the regression functions of the probability that the couple has a child on each wage ratio between wives and husbands. The estimated curves are shown in Figure 4, along with robust bias-corrected confidence intervals with 95% asymptotic coverage (Calonico et al., 2018, 2022). We include not only the case where the dependent variable is the percentage of couples who have ever had a child (Panel (a)), but also the case where the dependent variable is the number of children ever born (Panel (b)). Since the simple model considers only the extensive margin (whether a couple has children), the results in Panel (a) are suitable for testing the propositions of the simple model. Meanwhile, since the quantitative model in the next section considers the intensive margin (how many children a couple has), the results in Panel (b) fit the analysis there. Overall, the basic trend remains the same regardless of which is used as the dependent variable.

Let us first consider Proposition 1. Figure 2a in the Introduction shows that in both the US and Japan, women are the primary childcare providers in all wage ratio ranges analyzed. Thus, $\omega < \hat{\omega}$ holds for all wage ratios in our analysis. The result of Proposition 1 implies that the fertility curves for both the US and Japan should be downward sloping with respect to the wage ratio. The estimated curve in Figure 4a does not appear to be monotonically decreasing across the entire interval between 0 and 1. However, in regions close to 0 or 1, the sample size is small, and reliable estimation may not have been achieved. The wage share of wives in most households falls between 0.2 and 0.6. Within this range, the reliability of the estimates is relatively high, and the curve decreases monotonically. It demonstrates that the result of Proposition 1 is consistent with our data.[10]

[10] It is well known in the literature that the wife's wage and the husband's wage can have different effects

Figure 4: Regression Curve of Fertility on Relative Wages



(a) Prob. of Having Children

(b) Fertility Rate

Note: Local linear fits with IMSE-optimal bandwidths and robust bias-corrected confidence intervals with coverage error optimal bandwidths are shown. Source: CPS (2012, 2018) and JHPS (2012, 2019).

Proposition 2 suggests that a change in the social norm affects the fertility of different households in different directions. Relatively speaking, Japan can be regarded as an economy where social norms that require women to bear a greater burden of childcare are prevalent; that is, $\tilde{r}$ is large. Meanwhile, the US is an economy where such social norms are weak; that is, $\tilde{r}$ is small. Figure 4a suggests that the fertility probability is larger in the US for households in which the wife's wage is sufficiently higher than the husband's. Conversely, it is larger in Japan for households in which the wife's wage is sufficiently lower than the husband's. These estimation results are consistent with the results of Proposition 2.

In contrast to Proposition 2, which is about the difference in the level of the fertility curve between the US and Japan in Figure 4a, Proposition 3 refers to the difference in the slope of the fertility curve between the US and Japan. Proposition 3 implies that the slope of the fertility curve should be steeper in Japan with higher $\tilde{r}$ than in the US with lower $\tilde{r}$: In Japan, the group with higher wages of wives relative to their husbands has greater costs attributable to social norms. Hence, the decline in the number of children associated with an increase in wives' wage ratio should be larger than that in the US. The estimation

on fertility. Evidence suggests that the correlation between fertility and wife's wage is negative whether controlling for the husband's wage or not. Although evidence on the conditional correlation between fertility and husband's wage, holding the wife's wage constant, is mixed, some evidence like Blau and van der Klaauw (2008) find it positive. Given that women have been the primary childcare providers in most households, the result of Proposition 1 is also consistent with such evidence from previous studies.

results in Figure 4a support the validity of this prediction.

The previous analyses support the relevance of the prediction from our model. However, the problem of omitted variable bias may remain. To alleviate this problem, we also estimate a regression with controls. Specifically, we estimate

$$Y_i = \beta_0 + \beta_1 J_i + \beta_2 R_i + \beta_3 (J_i \times R_i) + Z_i'\gamma + \varepsilon_i, \tag{2.3}$$

where Y_i denotes the dummy for having children, J_i is the Japan dummy, R_i is the female relative wage, and Z_i includes the indicator if the family income is below the median family income of their country, the couple's ages and their interaction, the couple's education years and their interaction, and the indicators of full-time worker and their interaction, and 2012 dummy. Furthermore, we estimate two cases here: one restricted to couples whose children are six years old or older, and the other including couples whose children are younger than six years. This is to address the problem of potential selection bias. According to Kleven et al. (2024) and https://childpenaltyatlas.org, the wives' propensity to work after childbirth varies between the US and Japan, especially in the first five years after childbirth. If one country has a larger rate of women leaving the labor market when their children are under age 6 than the other one, a selection bias may arise when comparing only couples who remain in the labor market.

Table 2 summarizes the estimation results. Propositions 1, 2, and 3 imply that $\beta_2 < 0$ and $\beta_2 + \beta_3 < 0$, $\beta_1 > 0$, and $\beta_3 < 0$, respectively. The results in Table 2 confirm that these sign conditions are consistent with the data and are not rejected in any case. Note that since the baseline estimates for couples with children age 6 or older and those including couples with children under age 6 are similar, the selection bias described above should have a limited effect on the results in Table 2. In summary, our simple model reasonably captures the major patterns of fertility behavior in, and important differences between the US and Japan. In the next section, we build a quantitative model based on this simple model.

Although the above estimation of Equation (2.3) controls for several important factors, including the sample selection problem, one concern remains regarding reverse causality. Before proceeding to the next section, let us consider the problem of reverse causality. In Figure 4, we attempt to analyze how differences in the relative wages of married couples affect fertility. The downward curves are consistent with our theoretical model if we interpret them as being the result of couples with higher relative wages of the wife choosing to have fewer children. However, we cannot rule out the possibility that the direction of the effect is the opposite: The downward curves imply that the relative wages of wives declined for those households which chose to have more children. From such an opposite standpoint, the curves in Figure 4 can be interpreted as indicating that married women in Japan must have fewer children than in the US to maintain higher wages after childbirth; that is, the child penalty is greater in Japan than in the US. The analysis using event studies confirms that this is not the case in our data (See Appendix B for details).

Table 2: Difference in Response of Fertility to Relative Wage between the US and Japan

	Prob. of Birth Ever		
	(1)	(2)	(3)
$\hat{\beta}_1$ (Japan dummy J_i)	0.118	0.028	0.048
	(0.068)	(0.065)	(0.080)
$\hat{\beta}_2$ (Relative Wage R_i)	−0.142	−0.006	−0.098
	(0.055)	(0.053)	(0.061)
$\hat{\beta}_3$ (Interaction $J_i \times R_i$)	−0.372	−0.237	−0.302
	(0.169)	(0.168)	(0.213)
Include Controls Z_i		✓	✓
Exclude families with the youngest child aged ≤ 5			✓
N	2891	2829	2091

Note: The coefficient estimates from Equation (2.3) are shown, with the heteroskedasticity-robust standard errors in the parentheses. N indicates the sample size.

3 Quantitative Model

Here, we develop a quantitative model that we match to the data. Compared to the simple model described in the previous section, we: (i) introduce an overlapping generations structure, (ii) incorporate human capital, (iii) allow households to choose the number of children, and (iv) generalize the production function for childcare.

The model is an overlapping generations model where an individual lives for two periods, one as a child and one as an adult. A randomly matched couple comprises a woman f and man m, and characterized by corresponding human capital levels h_f and h_m, respectively. The utility of gender $j \in \{f, m\}$ is modified as follows:

$$u_j = \ln c_j + \varsigma_j \ln l_j + \gamma_j \ln n + \delta \ln h' - n \cdot \frac{\psi \left(r - \tilde{r}\right)^2}{2}.$$

Three changes have been made from the simple model. First, individuals derive utility from each child's human capital h', and δ is a parameter that expresses the degree of importance attached to the child's human capital. As explained later, a stochastic component exists in determining the child's human capital, but this component is assumed to be the same for all children in a household. There is no heterogeneity in human capital among the children of a given couple. Second, we allow two preference parameters, ς_j and γ_j, to differ between men and women. Third, the number of children is a non-negative real number, $n \in (0, \infty)$. With such a change, we adopt a functional form with diminishing marginal utility with respect to the number of children.

The male wage rate is determined by human capital: $w_m = h_m$. Meanwhile, to reflect the gender wage gap in the real labor market, we assume that the female wage rate is discounted by a gender wage gap factor $\chi \in (0, 1)$: $w_f = \chi h_f$. The welfare weight in the couple's problem is given by the wage ratio between the wife and husband: $\theta_f = w_f / \left(w_m + w_f\right) = \chi h_f / \left(h_m + \chi h_f\right)$.

Unlike the simple model, where parental time is the only input in childcare, we assume here that market childcare services are also available. The production function for childcare is given in form of constant elasticity of substitution (CES):

$$n = A\left[\alpha g^{\rho} + \left(x_f^{\epsilon} + x_m^{\epsilon}\right)^{\frac{\rho}{\epsilon}}\right]^{\frac{1}{\rho}}, \tag{3.1}$$

where g is the amount of market childcare services used, A and α respectively represent the efficiency of the childcare production function and the relative efficiency of market childcare compared to parental childcare, and $\rho \leq 1$ and $\epsilon \leq 1$ are parameters related to the elasticity of substitution between inputs: $1/(1-\rho)$ is the elasticity of substitution between market childcare services and parental childcare and $1/(1-\epsilon)$ is the elasticity of substitution between a husband and wife's childcare time. The childcare production function in the simple model is a special case with $\rho = 1$, $\epsilon = 1$, and $\alpha = 0$ in this general CES form.

Children's human capital is determined by three factors: basic human capital, private educational investment by parents, and stochastic component. Specifically, following de la Croix and Doepke (2003) and Kim et al. (2024), we assume the following functional form:

$$h' = \kappa\left(\iota + e^{\zeta}\right),$$

where ι is basic human capital, e denotes private educational investment by parents, κ represents an idiosyncratic ability shock that is unknown to parents when determining the amount of educational investment, and $\zeta \in (0,1)$ represents the return to private educational investment. The ability shock follows a log-normal distribution: $\ln\kappa \sim \mathcal{N}\left(0, \sigma_{\kappa}^2\right)$. We assume $\iota > 0$, ensuring a positive value for children's human capital even if parents do not invest in education.

Now, we define the couple's problem. The timing of events within a period is as follows. First, a man and a woman are matched, given the wage rates of the husband and wife. Then, the number of children is chosen, and children are born. Finally, before the children's abilities are revealed, husbands and wives determine their parenting time, their leisure time, their consumption, the amount of market childcare services, and the amount of educational investment in their children. Given the above assumptions about timing, the couple's decision problem can actually be formulated as a static problem in which all variables are simultaneously determined in a situation where the children's ability, κ, is not realized. The couple's problem is given by

$$\begin{aligned} V\left(h_f, h_m\right) = \max_{\{n,e,g,c_f,c_m,l_f,l_m,x_f,x_m\}} & \theta_f \mathbb{E}_{\kappa}\left[\ln c_f + \varsigma_f \ln l_f + \gamma_f \ln n + \delta \ln h' - n\frac{\psi\left(r-\tilde{r}\right)^2}{2}\right] \\ & + \theta_m \mathbb{E}_{\kappa}\left[\ln c_m + \varsigma_m \ln l_m + \gamma_m \ln n + \delta \ln h' - n\frac{\psi\left(r-\tilde{r}\right)^2}{2}\right], \end{aligned}$$

subject to

$$
\begin{aligned}
c_f + c_m + en + g &= \chi h_f \left(1 - l_f - x_f\right) + h_m \left(1 - l_m - x_m\right), \\
h' &= \kappa \left(\iota + e^{\zeta}\right), \\
n &= A \left[\alpha g^{\rho} + \left(x_f^{\epsilon} + x_m^{\epsilon}\right)^{\frac{\rho}{\epsilon}}\right]^{\frac{1}{\rho}}.
\end{aligned}
$$

Lastly, we define a stationary equilibrium and specify the set of conditions that the model must satisfy.

Definition 1. *A stationary equilibrium consists of couples' decision rules* $\{n, e, g, c_f, c_m, l_f, l_m, x_f, x_m\}$ *for each state and the distribution of human capital for men and women* $F\left(h_j\right)$ *in the next generation that satisfy the following conditions:*

- *Households' allocation rule solves the optimization problem defined above.*
- *The stationary distributions of human capital for men and women are fixed points:*

$$
\int_0^{\bar{h}} F\left(dh_f\right) = \int_0^{\bar{h}} F\left(dh_m\right) = 0.5 \times \frac{\sum_{\kappa} \pi_{\kappa} \int \int_{\left\{h \mid \mathfrak{h}\left(h_f, h_m\right) \leq \bar{h}\right\}} n\left(h_f, h_m\right) F\left(dh_f\right) F\left(dh_m\right)}{\int \int_h n\left(h_f, h_m\right) F\left(dh_f\right) F\left(dh_m\right)}, \tag{3.2}
$$

where π_{κ} *represents the probability of the ability shock and* $\mathfrak{h}\left(h_f, h_m\right)$ *is the human capital based on the decision rules* $n\left(h_f, h_m\right)$ *and* $x\left(h_f, h_m\right)$*. Equation (3.2) is derived assuming no gender difference in human capital.*

4 Calibration to the US Economy

This section outlines the calibration strategy, examines the model fit, and investigates the role of social norms in the model. We calibrate the stationary equilibrium to the US economy, and check whether the calibrated model can reproduce the average time allocation of wives and husbands and average fertility rate for each type of household when households are classified by the wife's wage ratio. Note that our analysis is limited to households with wife's wage ratios between 0.2 and 0.6. In the data used here, the proportion of couples in the US where the wife's wage ratio is equal to or less than 0.2 is 3.422%; meanwhile, the proportion of couples in Japan in which the wife's wage ratio is more than 0.6 is 6.906%. Thus, the range of wife's wage ratios that can be analyzed with reasonable accuracy for both countries is 0.2 to 0.6.

4.1 Externally Chosen Parameters

The gender wage gap factor χ is directly chosen from the data. In our samples from CPS (2012, 2018), the median of the wife's wage ratio is 0.833. Hence, we set $\chi = 0.833$. This value is quite close to the 2018 US gender wage gap of 0.811 based on OECD calculations.

Table 3: Calibration Results (US)

Parameter	Value	Target Moment	Data	Model
γ_m	0.369	Average childcare time (Husband)	0.079	0.075
γ_f	0.369	Average childcare time (Wife)	0.152	0.148
δ	0.997	Average share of education expenditures	0.022	0.022
ς_m	1.507	Average working hours (Husband)	0.399	0.398
ς_f	1.477	Average working hours (Wife)	0.333	0.336
ψ	3.058	Average fertility	1.771	1.746
$\tilde{r}$	0.667	Wife's share of childcare time at her wage share of 0.5	0.643	0.643
ι	2.617	Average income	1.000	1.010
ζ	0.644	Average share of education expenditures (Top household income quintile)	0.047	0.047
σ_k	0.411	Cumulative distribution of her wage share at 0.25	0.067	0.067
ϵ	0.550	Gap in Wife's share of childcare time (Wife's wage share of 0.2 vs. 0.6)	0.089	0.089
α	0.405	Average share of market childcare services	0.037	0.037
ρ	0.690	Average share of market childcare services (Top household income quintile)	0.062	0.061

In Equation (3.1), setting $x_f = 1$, $x_m = 0$, and $g = 0$ yields $n = A$. That is, A represents the number of children when the wife devotes all her time to childcare, the husband does not participate in childcare at all, and no market childcare services are used. From ATUS (2012, 2019), when limited to couples where the husband bears no childcare burden, the 95th percentile value for number of children is 3.8. Therefore, we set $A = 3.8$.

4.2 Internally Calibrated Parameters

Thirteen parameters remain to be determined. We pin down their values by minimizing the sum of squared percentage differences between the data and model moments. The moments used and calibration results are summarized in Table 3. Let us explain the relationship between the calibrated parameters and target moments. Obviously, no one-to-one correspondence exists between a given parameter and moment because each moment is affected by all parameters. However, some moments respond particularly strongly to changes in certain parameters. Thus, we group them into several categories based on such criteria and give an intuitive explanation.

First, the parameters in the utility function are γ_j, δ, ς_j, ψ, and $\tilde{r}$. The parameters γ_j, ς_j, and δ are weights for the quantity of children, leisure, and the quality of children in the utility function, respectively. They are pinned down by the average share of time spent on childcare and housework from ATUS (2012, 2019), average share of time spent on market labor from ATUS (2012, 2019), and average ratio of private education expenditures to income from the 2018 Consumer Expenditure Survey (CEX).[11] The parameter ψ represents the degree of concern for social norms: the higher the value of ψ, the greater the cost of children due to the presence of social norms, and thus, the more difficult it is for couples to have children. Thus, to capture such a characteristic of ψ as the cost of

[11]To compute the average ratio of private education expenditures to income, we focus on the second quarter of the CEX (2018), as it aligns most closely with the survey months of the JHPS, which is used to calibrate the model to Japan in the subsequent section.

having children, it is pinned down by the fertility rate. The targeted fertility rate is calculated based on responses to the question of how many children have been born so far in the CPS (2012, 2018); the computed value is 1.771. The parameter $\tilde{r}$ represents the ratio of the wife's childcare burden considered desirable in American society. To determine this parameter, we use the wife's average share of time spent on childcare and housework in couples where the wife's wage share is 0.5 (i.e., those with equal earning ability between spouses). The target value calculated from ATUS (2012, 2019) is 0.643.

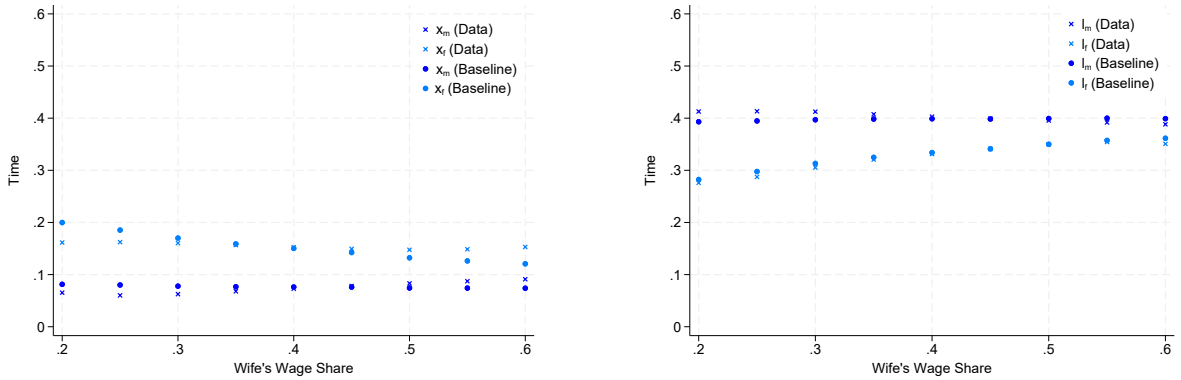
Second, the parameters related to the children's human capital function are ζ, ι, and σ_k. The return on private education investment, ζ, is calibrated using the average share of education expenditures for the top household income quintile. Since we include both the economy-wide average education spending share and the average education spending share of the top income quintile in our targets, the difference captures the response of education spending to changes in income, a proxy for the rate of return on education. According to the CEX (2018), the average education expenditure share of the top quintile is 4.69% while that of all households is 2.23%. The basic human capital, ι, is used as a normalization factor; it is set so that the average labor income of individuals in the model is close to 1.000. The parameter σ_k, which governs the dispersion of idiosyncratic ability shocks to children, is pinned down by the cumulative distribution of a wife's wage share at 0.25. According to CPS (2012, 2018), the cumulative distribution at this point is 0.067.

Lastly, the parameters related to the childcare production function are ϵ, α, and ρ. To calibrate the parameter governing the elasticity of substitution between a husband and wife, ϵ, we use the difference in the wife's average share of childcare time between couples with a wife's relative wage of 0.2 and 0.6 that is calculated from ATUS (2012, 2019). Note that the resulting value of ϵ is 0.550 and the elasticity of substitution implied by this value is 2.222, which is considerably smaller than the elasticity of substitution of 3.030 obtained in Knowles (2013). The most important reason for this divergence is likely that Knowles (2013) estimates the elasticity based on differences between two time points using the average male-female wage ratio and the average male-female housework time ratio in 1975 and 2003, whereas we use ATUS (2012, 2019) and CPS (2012, 2018) to estimate the elasticity based on cross-sectional household differences. If the estimate by Knowles (2013) also includes factors that reduce women's share of housework over time other than reduction in the gender wage gap, then it is a natural consequence that our estimate is considerably smaller than his. The relative efficiency of market childcare compared to parental childcare, α, and parameter related to the elasticity of substitution between market and parental childcare, ρ, are pinned down by the average share of household expenditures allocated to childcare services economy-wide and the average among households in the top income quintile from CEX (2018).

4.3 Model Fit

Next, we check the calibrated model's fit by comparing model moments to non-targeted data moments. Figures 5a and 5b show that the model matches the data extremely well

Figure 5: Childcare & Working Time by Wife's Wage Share (US)



(a) Childcare Time

(b) Working Time

Note: To calculate childcare and working time for her wage share in the baseline model, we average the corresponding shares for wife's wage share within the range $[\omega, \omega + 0.05)$. Source: ATUS (2012, 2019).

in terms of the household's time allocation for each wife's wage share.[12] It is noteworthy that while our target includes only the economy-wide average for each time use, the model fits well not only to the economy-wide average but also to all averages calculated for each wife's wage share.

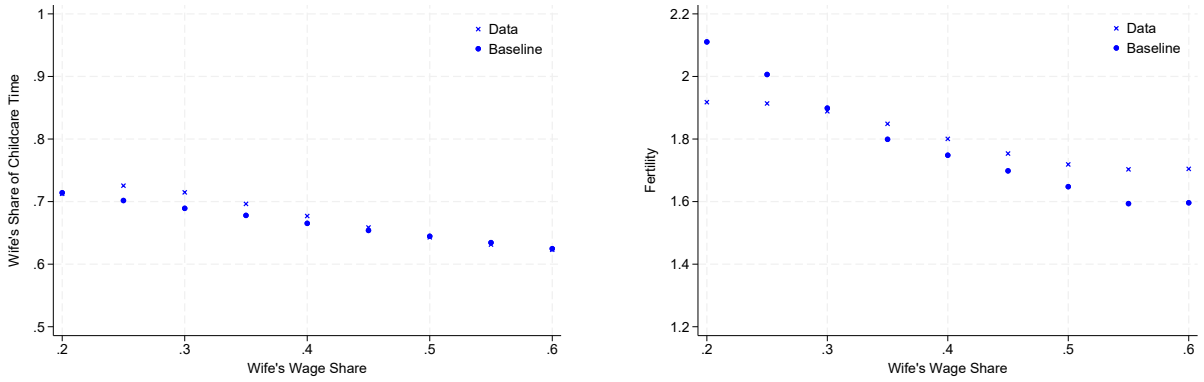
The other non-targeted moments used here to check the model fit are gender division of childcare and fertility, which are also the variables of primary interest in this study. Figure 6a shows that the gender division of childcare is reproduced quite well for all categories of the wife's wage share. As seen in Figure 5a, both levels before taking the ratio are well reproduced, so this result is not surprising, but it indicates that the main variables of this study are well reproduced in the baseline. Figure 6b illustrates fertility rates by a wife's wage ratio. The baseline captures well the overall level of fertility and the declining trend in fertility with increasing relative wages of wives. It overestimates the fertility rate for couples where the wife's wage ratio is less than 0.25 and underestimates that for couples where the ratio is 0.55 or higher, but it accurately reproduces the fertility rates for couples with other wife wage ratios. In the baseline, the proportions of households in which the wife's wage ratio is less than 0.25 and 0.55 or higher are 6.71% and 23.29%, respectively; in the data, they are respectively 6.69% and 21.74%. That is, groups with large discrepancies between the model's predictions and data account for a small percentage of the economy overall.

4.4 Role of Social Norms

To quantify the role of social norms in childcare sharing and fertility decisions, we conduct a counterfactual experiment in which social norms are absent, that is, we set $\psi = 0$

[12] To calculate each time use for households with a wife's wage share of ω in the baseline model, we average the corresponding time use for the wife's wage ratio within the range of $[\omega, \omega + 0.05)$.

Figure 6: Wife's Share of Childcare Time & Fertility by Wife's Wage Share (US)



(a) Wife's Share of Childcare Time

(b) Fertility

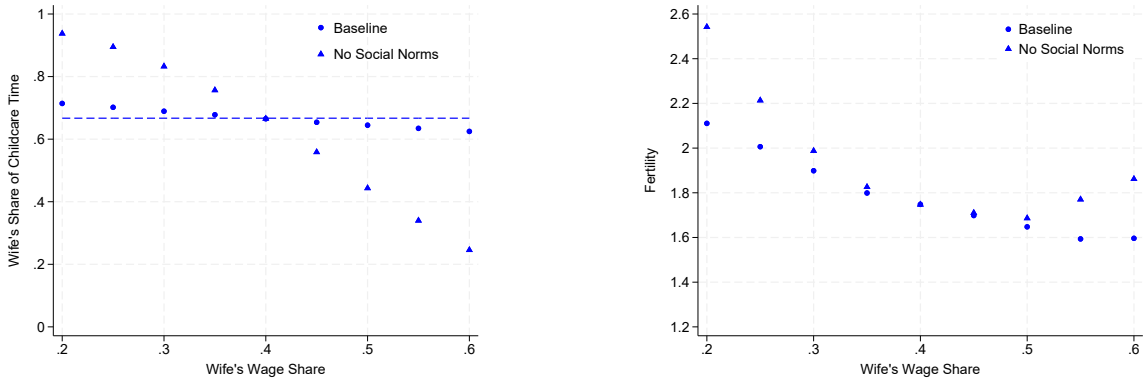
Note: To calculate the wife's share of childcare time and fertility for wife's wage share in the baseline model, we average the corresponding shares for wife's wage share within the range $[\omega, \omega + 0.05)$. The actual share and fertility are computed as in Figures 2a and 4b. Source: ATUS (2012, 2019) and CPS (2012, 2018).

while other parameters are fixed at the values obtained in the above calibration.

Figure 7a shows the wife's childcare time share for each wife's wage ratio in the baseline and counterfactual cases. Without social norms, couples can more efficiently allocate their childcare time based on the wages of the wife and husband in terms of income maximization. The change in the childcare time share in response to the spousal wage ratio is much more elastic in the absence of social norms. At the wage ratio of 0.2, a wife's share of childcare time is 0.938 (corresponding baseline value is 0.714), while at the wage ratio of 0.6, it drops to 0.246 (corresponding baseline value is 0.625). Note that the baseline and counterfactual childrearing time shares are equal around a relative wage of 0.4. This is because a couple with a relative wage of about 0.4 will choose $r^* = \tilde{r} = 0.667$, which is the socially desirable share estimated in the baseline, even in the absence of social norms. Thus, such couples choose the same burden share whether or not the social norm is present, and thus, there is no social norm-induced disutility.

Then, what happens to fertility? Figure 7b shows the fertility rate for each wife's relative ratio in the baseline and counterfactual cases. Our simulation indicates that without social norms, the fertility rate at $\omega \in [0.2, 0.6]$ increases from 1.771 to 1.848. This is because the presence of social norms reduces the benefits of having children because doing so entails disutility due to social norms, thus reducing the number of children. Even more interesting than the effect on overall fertility is that couples with different relative wages have different magnitudes of impact from social norms. Couples with a relative wage of about 0.4, where there is no disutility due to social norms, choose the same number of children with and without social norms. The further away from that wage ratio, the greater the disutility that the couple suffers, and thus, the greater the extent to which baseline fertility is lower than counterfactual fertility.

Figure 7: Wife's Share of Childcare Time and Fertility by Wife's Wage Share without Social Norms (US)



(a) Wife's Share of Childcare Time

(b) Fertility

Note: To calculate the wife's share of childcare time and fertility for a relative wage of x in the baseline model, we average the corresponding shares for relative wages within the range $[\omega, \omega + 0.05)$.

5 Calibration to the Japanese Economy and US-Japan Comparison

Next, we calibrate the stationary equilibrium to the Japanese economy, and examine how social norms differ between the US and Japan, and how such differences in social norms affect differences in fertility behavior between the two countries.

5.1 Calibration

Due to differences in the sets of data available between Japan and US, it is not possible to calibrate the Japanese economy using the same strategy applied to the US economy. One parameter (χ) is adjusted directly from the data, one parameter (α) is set exogenously, and three parameters (ϵ, σ_k, and ψ) are set to the same values as in the US baseline. The remaining nine parameters are recalibrated to the moments of the Japanese economy.

As in the calibration of the US economy, the gender wage gap factor is set directly from the data. Using the median of the wife-husband wage ratio in JHPS (2012, 2019), we set $\chi = 0.730$. This value is reasonably close to the 2018 gender wage gap of 0.765 for Japan as calculated by the OECD. The efficiency of market childcare compared to parental childcare in Japan is set at half the value in US: $\alpha = 0.202$. Based on the fact that Japanese households employ nannies and similar services less than US households, we set the Japanese value at half the US value. It cannot be denied that this setting is arbitrary. However, similar results are obtained whether we adopt a value close to zero or one closer to the US value. The parameter related to the elasticity of substitution between spouses in childcare, ϵ, the dispersion of idiosyncratic ability shocks to children, σ_k, and the degree of concern for social norms, ψ, are set to their corresponding US values.

Table 4: Calibration Results (Japan)

Parameter	Value	Target Moment	Data	Model
γ_m	0.558	Average childcare time (Husband)	0.060	0.059
γ_f	0.575	Average childcare time (Wife)	0.340	0.341
δ	1.209	Average income share spent on education expenditures	0.065	0.065
ς_m	1.313	Average working hours (Husband)	0.468	0.465
ς_f	1.079	Average working hours (Wife)	0.269	0.269
$\tilde{r}$	0.914	Wife's share of childcare time at her wage share of 0.5	0.816	0.811
ι	2.055	Average relative wage to the US	0.801	0.805
ζ	0.481	Average share of education expenditures (Top household income quintile)	0.085	0.085
A	2.451	Average Fertility	1.512	1.517

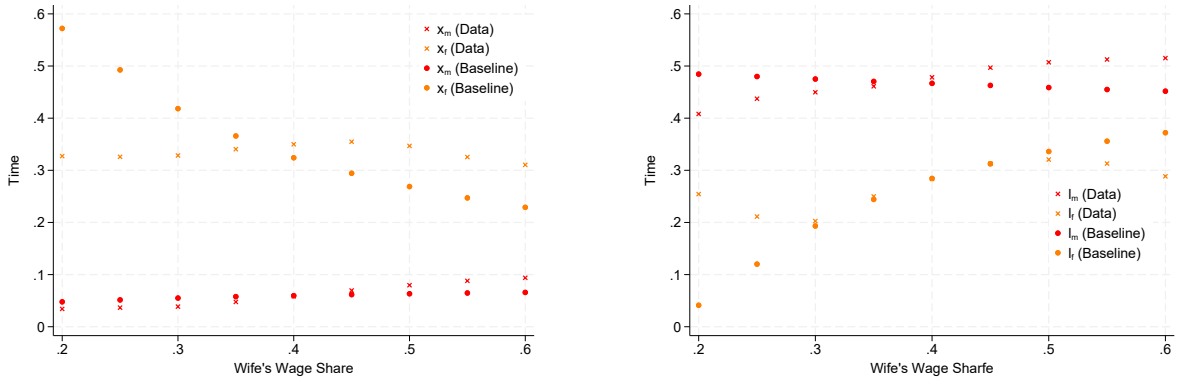
We recalibrate the remaining nine parameters, γ_f, γ_m, δ, ς_f, ς_m, $\tilde{r}$, ι, ζ, and A to the Japanese economy. The data moments are: the average childcare time for wives, average childcare time for husbands, average ratio of education expenditures to household income, average working hours for wives, average working hours for husbands, share of wife's childcare time for couples where the wife's wage ratio is 0.5, average wage in Japan when the average wage in the US is normalized to 1, average share of education expenditures for the top household income quintile, and fetility rate. We use the 2016 Survey on Time Use and Leisure Activities to calculate the ratio of leisure time. The JHPS (2012, 2019) is used to calculate the other data moments.

The results are summarized in Table 4. Let us compare the US and Japan on several parameters. The weight on leisure, ς_j, is larger in the US, reflecting the fact that leisure time, calculated as total time minus childcare and housework time and market labor time, is longer in the US. In the US, there is little difference between men and women regarding preference parameters. In Japan, however, while preferences for children show little gender difference, preferences for leisure time are estimated to be smaller for women than for men. This stems from the fact that women in Japan have less leisure time than men. The return on educational investment, ζ, is higher in the US, reflecting the fact that in Japan, the average education spending ratio of quintile households is only about 30% higher than the average of all households; in the US, it is about 120% higher. The basic human capital, ι, is larger in the US, reflecting that the average wage in the US is higher than in Japan. Finally, the most notable feature is the difference between the US and Japan in $\tilde{r}$, the wives' share of childcare and housework considered ideal in that society. The estimated value of $\tilde{r}$ in Japan is 0.914, much higher than the US estimated value of 0.667 and close to 1. This indicates the social norm in Japan that almost all childcare and housework is the work of women.

5.2 Model Fit

How well does the model with the recalibrated parameters capture the Japanese economy? Figure 8a and 8b illustrate households' time allocation for each wife's wage share. The model performs reasonably well overall, but it fails to replicate the wife's time allocation when her wage share is low. This shortcoming may be partly explained by the

Figure 8: Childcare & Working Time by Wife's Wage Share (JP)



(a) Childcare Time

(b) Working Time

Note: To calculate childcare and working time for her wage share in the baseline model, we average the corresponding shares for wife's wage share within the range $[\omega, \omega + 0.05)$. Source: JHPS (2012, 2019).

limited flexibility of working hours in the Japanese labor market compared to the US labor market. Apart from this limitation, the model captures the data reasonably well.

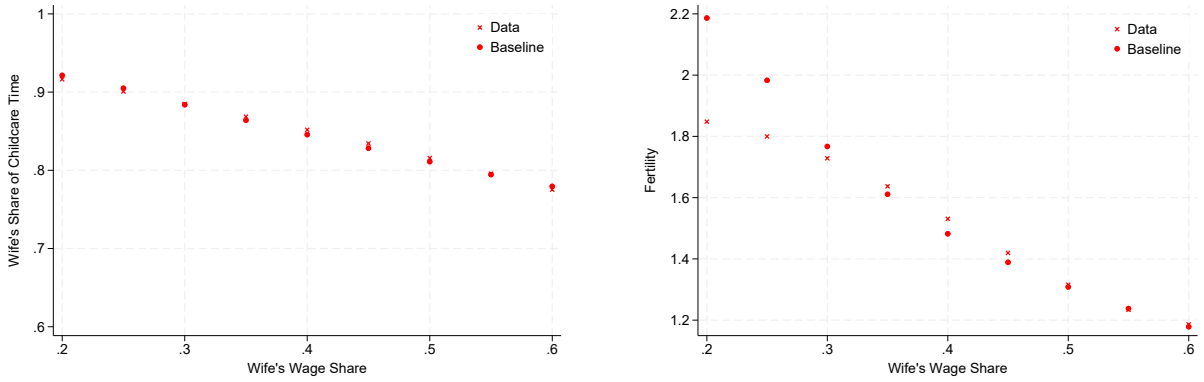
Figure 9a compares the share of a wife's childcare time by her wage share in the data and in the baseline model. It can be seen that the model replicates the data quite well. The next non-targeted moment is the fertility rate by wife's wage share, the fit of which is shown in Figure 9b. Although there is a non-negligible fertility gap between the data and baseline model for couples with wives' wage share below 0.3, our model accurately reproduces the fertility of the remaining couples, which accounts for 80.11% of the data. Notably, our model correctly captures the steep decline in fertility associated with the increase in the wife's wage share in Japan relative to the US.

5.3 Relationship between Social Norms and Differences in Fertility

Our baseline model successfully replicates the marital childcare sharing and fertility by wife's wage ratio for both the US and Japan. An interesting question then arises: To what extent can the differences in fertility behavior between the two countries be explained by the differences in social norms? To address this question, we conduct a counterfactual experiment in which US social norms are applied to the Japanese economy. In the model calibrated to the Japanese economy in the previous subsection, we change only $\tilde{r}$ to the value obtained by calibrating to the US economy (i.e., $\tilde{r} = 0.667$), while keeping the other parameters unchanged.

Figure 10a plots the wife's parenting time share by wife's wage ratio under two different social norms. As expected, when the US social norm is applied, the wife's parenting time share is remarkably close to that of the US economy. The wife's parenting time share for couples with a wife's wage share of 0.5 is 0.631 (0.645 in the US and 0.811 in the Japanese baseline), and the difference in parenting time shares between couples with a wife's wage share of 0.2 and 0.6 is 0.147 (0.089 in the US and 0.142 in the Japanese

Figure 9: Wife's Share of Childcare Time & Fertility by Wife's Wage Share (Japan)



(a) Wife's Share of Childcare Time

(b) Fertility

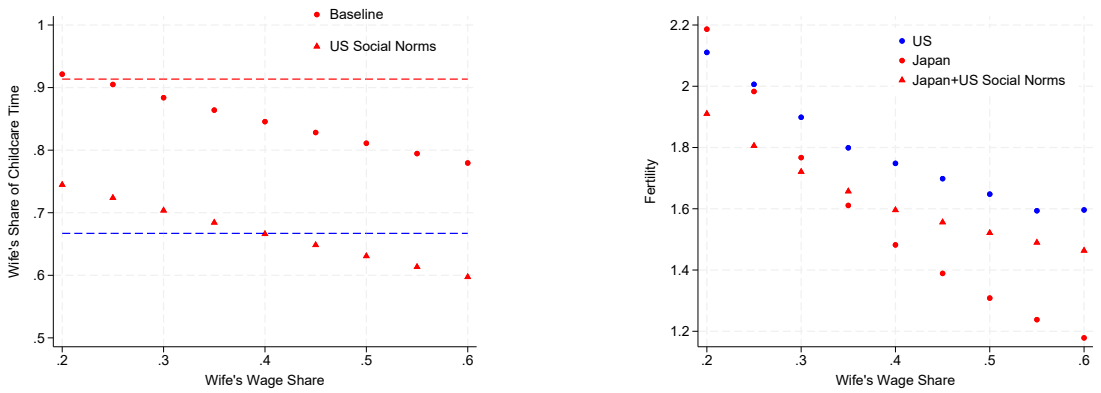
Note: To calculate the wife's share of childcare time and fertility for a wage share of ω in the baseline model, we average the corresponding shares for wage shares within the range $[\omega, \omega + 0.05)$. The actual share and fertility are computed as in Figures 2a and 4b. Source: JHPS (2012, 2019).

baseline).

Then, what happens to fertility? Our simulations show that the fertility rate at $\omega \in [0.2, 0.6]$ increases from 1.564 to 1.642. As shown in Figure 10b, in the Japanese economy, when social norms change to those of the US, the average fertility rate decreases for couples in which the wife's wage share is less than 0.35, but increases for all other couples. Since couples whose wives earn a wage share of 0.35 or more account for about 68.78% of the Japanese economy, the number of couples who have more children is much larger and the fertility rate substantially increases. The mechanism behind this result is essentially the same as that for the simple model in Proposition 2: as a result of the change in social norms, the fertility rate of couples whose chosen division of childcare moves closer to the efficient division of childcare in terms of income maximization rises, while the fertility rate of couples whose chosen division of childcare moves farther away from it declines. Replaced by the US estimated $\tilde{r}$, for couples whose wives' wage share is less than 0.35, the fertility rate falls as the discrepancy from the efficient share widens in the opposite direction, too low, but for couples with wives' wage share above 0.35, fertility rises as $\tilde{r}$ approaches their efficient wife's share. Furthermore, applying US social norms, the decline in fertility associated with an increase in the wife's wage share becomes smaller. This is exactly what Proposition 3 of the simple model states. Consequently, the higher the wife's wage share, the larger the increase in fertility due to the social norm being the same as in the US. For a couple with a wife's wage share of 0.6, the fertility rate increases by 0.285 from 1.178 to 1.463.

Finally, to quantify how much social norms contribute to the fertility differential between the US and Japan, let us calculate the proportion of the fertility gap between the two countries which is attributable to social norms. The proportion is defined as the fertility differential between the two social norms in Japan, divided by the fertility differential between the US and Japan. Thus, the value obtained is (1.642-1.564)/(1.771-1.564) = 0.377.

Figure 10: Wife's Share of Childcare Time and Fertility by Wife's Wage Share under the US Social Norms



(a) Wife's Share of Childcare Time

(b) Fertility

Note: To calculate the wife's share of childcare time and fertility for a wage share of ω in the baseline model, we average the corresponding shares for wage shares within the range $[\omega, \omega + 0.05)$. Source: JHPS (2012, 2019).

Surprisingly, about 38% of the difference in fertility rates between the two countries can be explained by the differences in social norms.

5.4 Social Norms vs. Educational Investment and Market Childcare

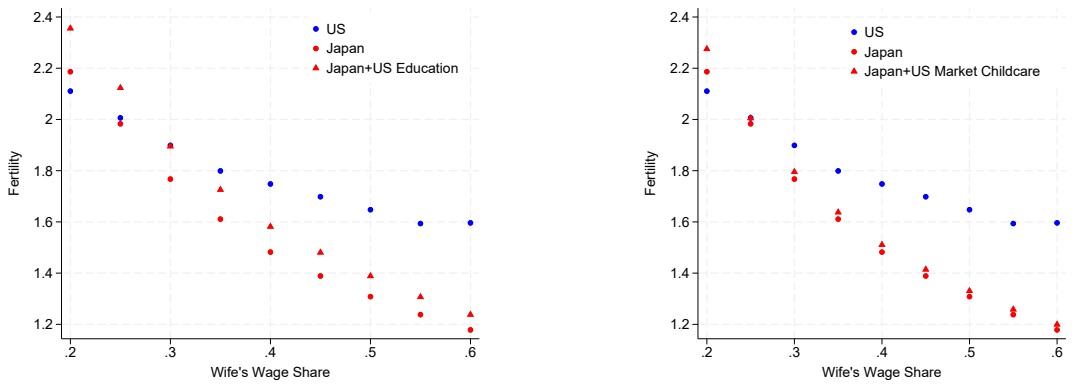
To relatively assess the impact of social norms on the difference in the fertility rates between Japan and the US, this subsection conducts counterfactual experiments examining what would happen if factors other than social norms, considered important for explaining the US-Japan fertility gap, were identical across both countries.

First, consider the impact of educational investment in children on the US-Japn fertility gap. In the Japanese baseline model, while keeping all other parameters unchanged, we set the preference parameter for children's human capital, δ, and the parameter for the return on educational investment, ζ, to equal those of the US. That is, we change the value of δ from 1.209 to 0.997 and the value of ζ from 0.481 to 0.644. Figure 11a illustrates the result. As education spending decreases, the fertility rate increases for all ω, and as a result, the overall fertility rate calculated for $\omega \in [0.2, 0.6]$ increases from 1.564 to 1.674.

Second, consider the impact of market childcare on the US-Japan fertility gap. In the Japanese baseline model, while all other parameters remain unchanged, the relative efficiency of market childcare compared to parental childcare, α, is set to equal that of the US. That is, we change the value of α from 0.202 to 0.405. Figure 11b illustrates the result. Market childcare services become convenient, leading to an increase in the fertility rate for all ω, and consequently, the overall fertility rate calculated for $\omega \in [0.2, 0.6]$ increases from 1.564 to 1.612.

Figures 10b, 11a, and 11b indicate that social norms, educational investment, and market childcare have each a significant impact on the fertility gap between the US and Japan.

Figure 11: Fertility by Wife's Wage Share with the US Education Expenditures & US Market Childcare



(a) US Education Expenditures

(b) US Market Childcare

Note: To calculate the wife's share of childcare time and fertility for a wage share of ω in the baseline model, we average the corresponding shares for wage shares within the range $[\omega, \omega + 0.05)$. Source: JHPS (2012, 2019).

However, a notable difference between the effects of social norms and other factors is that while the other factors affect only the overall level, social norms affect the slope of the curve as well as the overall level. This is particularly evident in the impact on the fertility gap between Japan and the US in households with a high wife's wage ratio. Focusing on households with $\omega = 0.6$, let us compare the results of the counterfactual analysis. The fertility rate for Japanese couples with $\omega = 0.6$ is 1.178, while the corresponding rate for American couples is 1.596. The Japanese fertility rate for couples with $\omega = 0.6$ becomes 1.463 when social norms change, 1.238 when education investment-related parameters change, and 1.199 when the efficiency of market childcare changes. Therefore, changing social norms narrows the gap by approximately 68%, changing education investment narrows it by approximately 14%, and changing market childcare narrows it by approximately 5%. While all the three factors have significant effects on the overall fertility rate, social norms are particularly important in households with a high wife's wage share.

6 Conclusion

Among developed economies facing declining fertility rates, the US belongs to a group with relatively high fertility rates, while Japan belongs to a group with relatively low fertility rates. Our microdata analysis of the US and Japan reveals that most of the difference in fertility rates between the two countries stems from differences in fertility rates among households where the wife's wages are relatively high compared to the husband's. This study demonstrates that a model incorporating social norms regarding the division of childcare burden between men and women can successfully explain such differences in fertility patterns between the US and Japan. Furthermore, the counterfactual analysis

conducted based on our model reveals that the majority of the difference in fertility rates between Japan and the US in households where the wife's wages are relatively high can be explained by differences in social norms.

Our analysis suggests several future research directions. First, the model can be extended to include own educational decisions. In our model, the costs incurred from social norms depend on married couples' wage shares. In Japan, where women are required to bear the greater burden of child rearing, couples with high wage shares of women suffer relatively greater costs. In the US, where this is not the case, couples with low wage shares of women suffer relatively greater costs than in Japan. Thus, in educational decisions prior to marriage formation, there is an incentive for men to get more education and women to get less education in Japan, so as to avoid couples in which women have higher wage shares. While most OECD countries, including the US, have higher college enrollment rates for women, Japan is one of the few exceptions where men have higher college enrollment rates. Endogenizing educational decision making based on our model may explain such cross-country differences in gender education differentials.

Second, our analysis can be applied to understand the effects of long-run changes in social norms. This study focuses on the behavior of married couples by wife-husband wage ratio using microdata, and since no corresponding long-term data are available, we are limited to a static US-Japan comparison. However, social norms do change in the long run, and women's labor market status also changes in the long run. It would be interesting to analyze the impact of the interaction between such social norms and women's labor market status on changes in the historical gender division of labor and fertility.

Third and last, it would also be an interesting attempt to extend our model to be able to analyze match formation and extensive margin of fertility decision. In this study, we focus on the behavior of married couples, with the distribution of married couples differentiated by each wife's wage share given exogenously. However, social norms, such as those analyzed in this study, should also affect the distribution of wives' wage shares through their influence on the type of partner with whom they marry. In addition, while this study focuses on differences in the fertility rates of couples having children in the US and Japan, the proportion of men and women having children is also important for differences in the total fertility rates between the two countries. Thus, a more comprehensive comparison of fertility behavior requires the consideration of the extensive margin of fertility decision.

A Couple's maximization

The Lagrangian is given by

$$\begin{aligned}\mathcal{L} &= \theta_f\left(\ln c_f + \varsigma \ln l_f\right) + \theta_m\left(\ln c_m + \varsigma \ln l_m\right) + \gamma - \frac{\psi\left(r-\tilde{r}\right)^2}{2} \\ &\quad + \lambda\left\{w_f\left(1 - l_f - r\phi\right) + w_m\left[1 - l_m - (1-r)\phi\right] - c_f - c_m\right\}.\end{aligned}$$

In the case of an interior solution, first-order conditions are

$$\frac{\theta_f}{c_f} = \lambda, \tag{A.1}$$

$$\frac{\theta_f \varsigma}{l_f} = \lambda w_f, \tag{A.2}$$

$$\frac{\theta_m}{c_m} = \lambda, \tag{A.3}$$

$$\frac{\theta_m \varsigma}{l_m} = \lambda w_m, \tag{A.4}$$

$$-\psi\left(r - \tilde{r}\right) - \lambda\phi\left(w_f - w_m\right) = 0, \tag{A.5}$$

and

$$w_f\left(1 - l_f - r\phi\right) + w_m\left[1 - l_m - (1-r)\phi\right] - c_f - c_m = 0. \tag{A.6}$$

Substituting (A.1), (A.2), (A.3), and (A.4) into (A.6), we obtain

$$w_f\left(1 - r\phi\right) + w_m\left[1 - (1-r)\phi\right] = \frac{1+\varsigma}{\lambda}.$$

Using this and (A.5), we obtain (2.2).

The associated consumption and leisure are as follows:

$$c_f = \frac{\theta_f}{1+\varsigma}\left\{w_f\left(1 - r\phi\right) + w_m\left[1 - (1-r)\phi\right]\right\},$$

$$c_m = \frac{\theta_m}{1+\varsigma}\left\{w_f\left(1 - r\phi\right) + w_m\left[1 - (1-r)\phi\right]\right\},$$

$$l_f = \frac{\theta_f \varsigma}{w_f\left(1+\varsigma\right)}\left\{w_f\left(1 - r\phi\right) + w_m\left[1 - (1-r)\phi\right]\right\},$$

and

$$l_m = \frac{\theta_m \varsigma}{w_m\left(1+\varsigma\right)}\left\{w_f\left(1 - r\phi\right) + w_m\left[1 - (1-r)\phi\right]\right\}.$$

B Event Study

To consider the possibility of reverse causality in Figure 4, we test whether the negative impact of the birth of a child on the mother's subsequent wages is greater in Japan than in the US. To perform such a test, we use the event study design proposed by Kleven et al. (2019). Kleven et al. (2019) analyze the impact of child birth on women's subsequent wages in Denmark, and the same methodology is applied for the US by Cortés and Pan (2023). We apply the same methodology for Japan and compare the magnitude of the impact between Japan and the US by comparing our result with that of Cortés and Pan (2023).

We again use the JHPS data. Unlike the previous sections, we use panel data from 2004 to 2020. The sample is restricted to households that meet the following three conditions: (1) both husband and wife are between the ages of 20 and 45 at the birth of their first child; (2) information is available for both the year before and the year of birth; and (3) they have participated in the survey for at least four years.

As in Cortés and Pan (2023), we use individuals who are working in a given year and estimate the impact of childbirth on wage rates conditional on working. Specifically, we estimate the following event-study regression:

$$\ln W_{ist}^{j} = \sum_{h\in\{-5,\ldots,-2,0,\ldots,10\}} \beta_h^j \mathbf{1}\{t=h\} + \sum_{k=22}^{55} \gamma_k^j \mathbf{1}\{A_{is}=k\} + \sum_{l=2004}^{2020} \delta_l^j \mathbf{1}\{s=l\} + \varepsilon_{ist}^j, \tag{B.1}$$

where W_{ist}^{j} represents the wage rate of individual i of gender j in year s; t denotes the number of years since birth, for example, -1 indicates one year before birth; and A_{is} represents the age of i in year s.[13] Note that β_{-1}^{j} is normalized to 0: β_h^j's capture percentage changes compared to time $t=-1$.

Figure 12 shows the results. It can be seen that the decline in women's wages after the birth of a child in Japan is not as great as in the US, that is, the child penalty for Japanese women is not as great as that for US women. Therefore, the analysis of the event study does not support that the results in Figure 4 are caused by reverse causality.

References

Albanesi, S. and Olivetti, C. (2009). Home Production, Market Production and the Gender Wage Gap: Incentives and Expectations. *Review of Economic Dynamics*, 12(1):80–107.

Beach, B. and Hanlon, W. W. (2022). Culture and the Historical Fertility Transition. *The Review of Economic Studies*, 90(4):1669–1700.

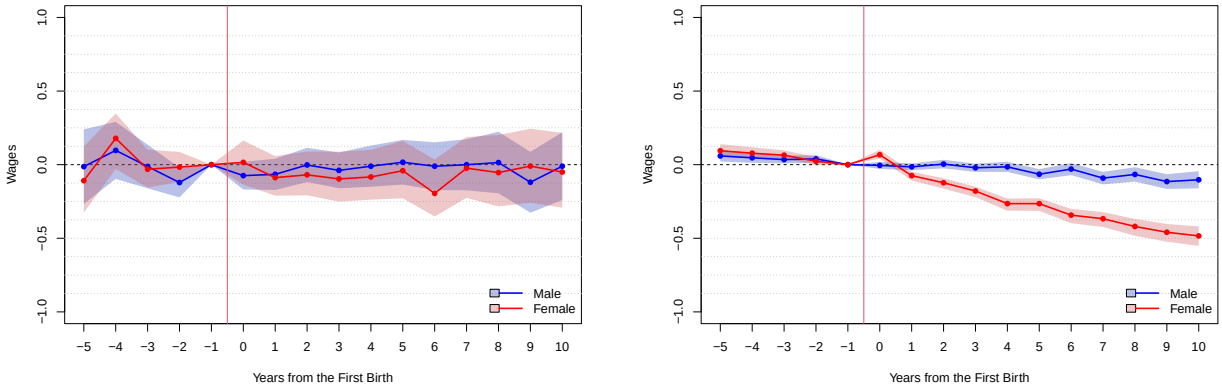
Bertrand, M., Cortes, P., Olivetti, C., and Pan, J. (2021). Social Norms, Labour Market Opportunities, and the Marriage Gap Between Skilled and Unskilled Women. *The Review of Economic Studies*, 88(4):1936–1978.

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

Blau, D. M. and van der Klaauw, W. (2008). A Demographic Analysis of the Family Structure Experiences of Children in the United States. *Review of Economics of the Household*, 6(3):193–221.

[13] Kleven et al. (2019) propose an alternative measure of percentage effect to avoid the "log of zero" problem. However, since we focus on working individuals, we do not face such a problem and thus simply take the logarithm for the dependent variable.

Figure 12: Effect of Birth on Wage Rates



(a) Japan **(b)** US

Note: The coefficients β_h^j are shown. The shaded area indicates the 95% (pointwise) heteroskedasticity-robust standard errors clustered at individual levels. (b) shows Cortés and Pan (2023, Figure 2D) using (B.1) and their provided replication data. Source: JHPS (2004-2020).

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2018). On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference. *Journal of the American Statistical Association*, 113(522):767–779.

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2022). Coverage Error Optimal Confidence Intervals for Local Polynomial Regression. *Bernoulli*, 28(4):2998–3022.

Cortés, P. and Pan, J. (2023). Children and the Remaining Gender Gaps in the Labor Market. *Journal of Economic Literature*, 61(4):1359–1409.

de la Croix, D. and Doepke, M. (2003). Inequality and Growth: Why Differential Fertility Matters. *American Economic Review*, 93(4):1091–1113.

de Laat, J. and Sevilla-Sanz, A. (2006). Working women, men's home time and lowest-low fertility. *ISER Working Paper Series, No. 2006-23, University of Essex, Institute for Social and Economic Research (ISER)*.

de Silva, T. and Tenreyro, S. (2020). The Fall in Global Fertility: A Quantitative Model. *American Economic Journal: Macroeconomics*, 12(3):77–109.

Doepke, M., Hannusch, A., Kindermann, F., and Tertilt, M. (2023). Chapter 4 - The economics of fertility: a new era. In Lundberg, S. and Voena, A., editors, *Handbook of the Economics of the Family*, volume 1 of *Handbook of the Economics of the Family*, pages 151–254. North-Holland.

Doepke, M. and Kindermann, F. (2019). Bargaining over Babies: Theory, Evidence, and Policy Implications. *American Economic Review*, 109(9):3264–3306.

Fernández, R. and Fogli, A. (2006). Fertility: The Role of Culture and Family Experience. *Journal of the European Economic Association*, 4(2-3):552–561.

Fernández, R. and Fogli, A. (2009). Culture: An Empirical Investigation of Beliefs, Work, and Fertility. *American Economic Journal: Macroeconomics*, 1(1):146–77.

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

Galor, O. and Weil, D. N. (1996). The Gender Gap, Fertility, and Growth. *The American Economic Review*, 86(3):374–387.

Hara, H. and Rodriguez-Planas, N. (2025). Long-Term Consequences of Teaching Gender Roles: Evidence from Desegregating Industrial Arts and Home Economics in Japan. *Journal of Labor Economics*.

ISSP Research Group (2016). International Social Survey Programme: Family and Changing Gender Roles IV - ISSP 2012. ZA5900 Data file Version 4.0.0.

Jones, L. E., Schoonbroodt, A., and Tertilt, M. (2010). 2. Fertility Theories: Can They Explain the Negative Fertility- Income Relationship? In Shoven, J. B., editor, *Demography and the Economy*, pages 43–106. University of Chicago Press, Chicago.

Kim, S., Tertilt, M., and Yum, M. (2024). Status Externalities in Education and Low Birth Rates in Korea. *American Economic Review*, 114(6):1576–1611.

Kimura, M. and Yasui, D. (2010). The Galor–Weil Gender-Gap Model Revisited: From Home to Market. *Journal of Economic Growth*, 15(4):323–351.

Kleven, H., Landais, C., and Leite-Mariante, G. (2024). The Child Penalty Atlas. *The Review of Economic Studies*, page rdae104.

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

Knowles, J. A. (2013). Why are Married Men Working So Much? An Aggregate Analysis of Intra-Household Bargaining and Labour Supply. *The Review of Economic Studies*, 80(3):1055–1085.

Kuehnle, D., Oberfichtner, M., and Ostermann, K. (2021). Revisiting Gender Identity and Relative Income within Households: A Cautionary Tale on the Potential Pitfalls of Density Estimators. *Journal of Applied Econometrics*, 36(7):1065–1073.

Myong, S., Park, J., and Yi, J. (2021). Social Norms and Fertility. *Journal of the European Economic Association*, 19(5):2429–2466.

Online Appendix for "The Role of Social Norms in Gender Division of Childcare and Fertility: A Quantitative Analysis of US-Japan Comparison"

Tomoaki Kotera, Yuta Okamoto, and Daishin Yasui

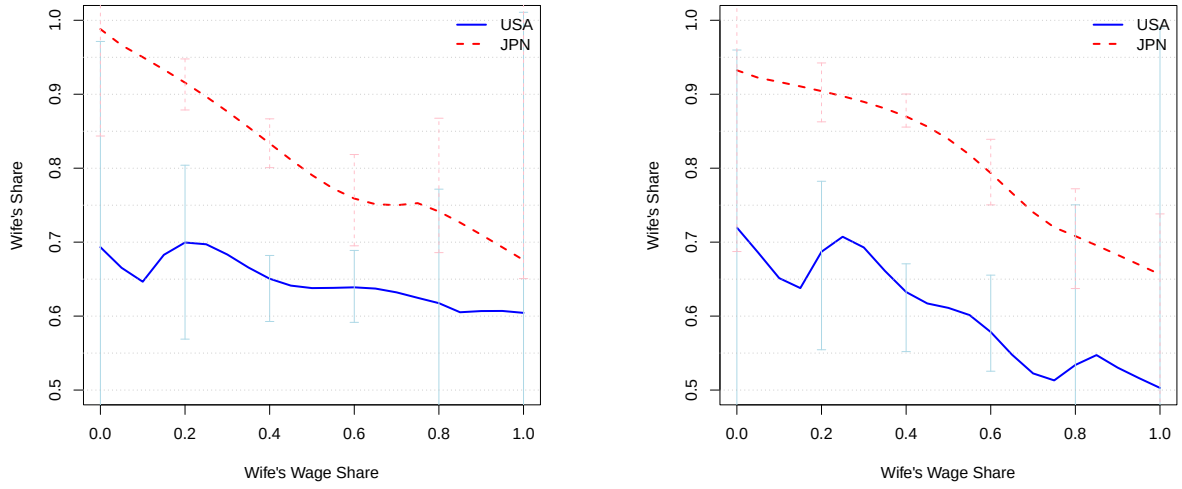
September 14, 2025

S1 Sensitivity Checks

S1.1 Figure 2a

In Figure S1, we show the female share of childcare and other housework burden by relative wage among families with children. Both exhibit a similar pattern and are compatible with Figure 2a.

Figure S1: Female's Share of Domestic Work & Relative Wage



(a) Share of Childcare by Relative Wage

(b) Share of Other Housework by Relative Wage

Note: Source: ATUS (2012, 2019) and JHPS (2012, 2019).

S1.2 Figure 3

In constructing the attitude measure in Figure 3, we assign values of 1, 0.5, and 0 to respondents who said mothers should work full time, part time, and stay home, respectively. However, there is no reasonable basis for assigning 0.5 to part-time, which may be

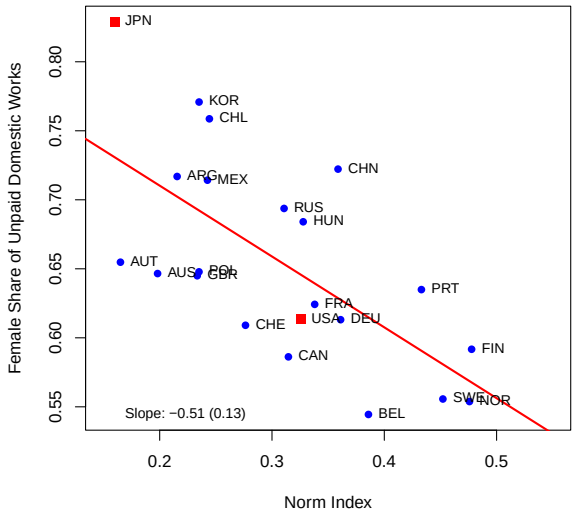
arbitrary. We therefore check to see if the pattern in Figure 3 remains the same when we change the value assigned to part-time. The results are provided in Figure S2. Regardless of whether we assign 1/3 or 2/3 to part-time, the pattern in Figure 3 remains almost the same.

S2 Definitions of Childcare and Housework Activities

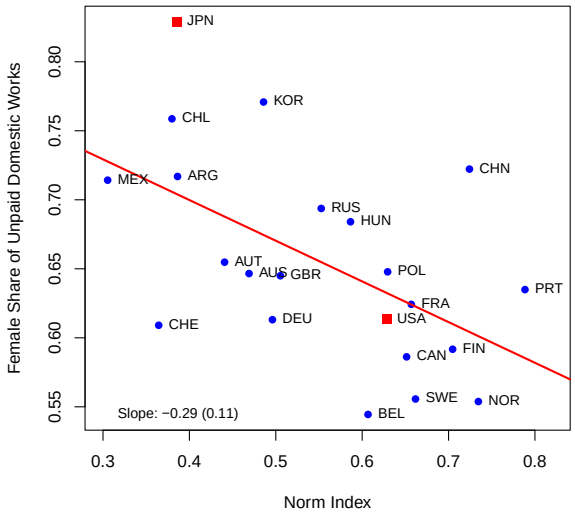
We utilized the ATUS data to construct variables on time use for childcare and housework. The definitions of these are provided below. Childcare is defined as the sum of the ATUS codes 0301 (caring for & helping household Children), 0302 (activities related to household children's education), 0303 (activities related to household children's health), and 180381 (travel related to caring for and helping household children). Housework is defined as the sum of the codes 02 (household activities), 03 (caring for & helping household members), 04 (caring for & helping nonhousehold members), 07 (consumer purchases), 08 (professional & personal care services), 09 (household services), 10 (government services & civic obligations), and 18 (traveling), excluding 0205 (lawn, garden, and houseplants) and 0206 (animals and pets), as well as household and non-household childcare activities and their associated travel.

For Japanese JHPS data, childcare and housework are simply defined as the time spent on the "childcare" and "housework," respectively.

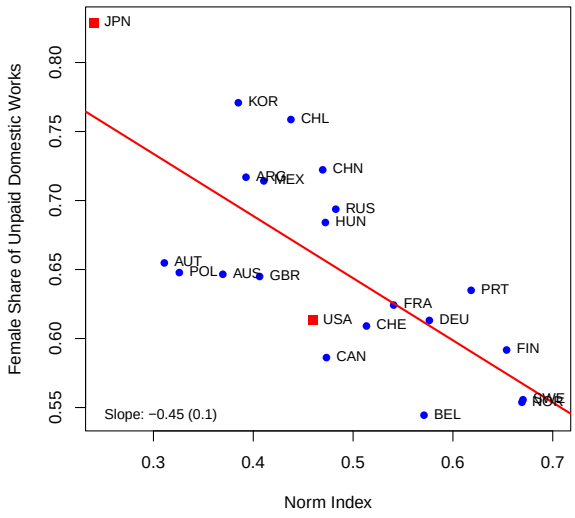
Figure S2: Social Attitudes and Domestic Work Division



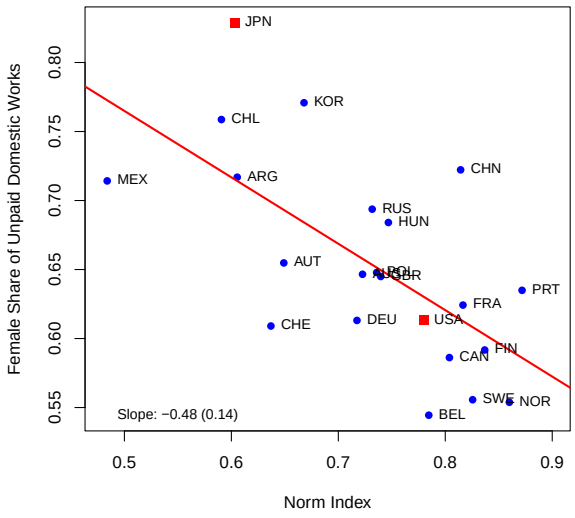
(a) The Youngest Kids under School Age



(b) After the Youngest Kids Starts School



(c) The Youngest Kids under School Age



(d) After the Youngest Kids Starts School

Note: In the upper panel, part-time is assigned a value of 1/3, while in the lower panel, a value of 2/3 is assigned to part-time. Source: ISSP (2012).