Explaining the Fall in Female Labor Supply in Urban China

Pengzhan Qian*

Queen Mary University of London

February, 2022

Abstract

Female labor force participation in urban China has fallen by 10% from the mid-1990s to the 2010s. This unusual trend is driven by a dramatic fall in participation among women with high-school degrees or lower, notwithstanding a substantial growth in their real incomes. This paper studies the life cycle labor supply of non-college women in cohorts 1950, 1960, and 1970. It features a household, life-cycle model with endogenous female participation to investigate which channels account for the fall in participation among the younger cohorts. The main finding is that family-related channels are equally important as income-related channels. Through income channels, a widening gender pay gap can explain 50% of the decline in the participation rate in low-educated women and 80% of that in medium-educated women. For family channels, increased childcare cost explains most of the remaining decline in labor supply. Increased assortative marriage has heterogeneous effects on different education levels. Finally, the difference between the 1950 and 1970 cohorts explains around 40% of the declined participation between 1989 and 2009.

Keywords: Female labor supply; gender pay gap; household; China.

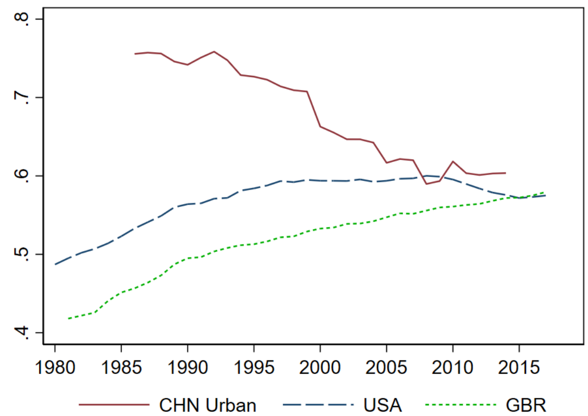
JEL Code: J12, J13, J21, J24, J31

*Acknowledgement: I thank the guide of my supervisor Giulio Fella and Yu Zheng. I thank the suggestions of Tim Lee, Anna Raute, Vincent Sterk, Wenchao Jin, and all participants and comments from QMUL internal seminars, DSE Summer School 2021, DIAL Final Conference, Austrian Labor Market Workshop, and EWMES 2022.

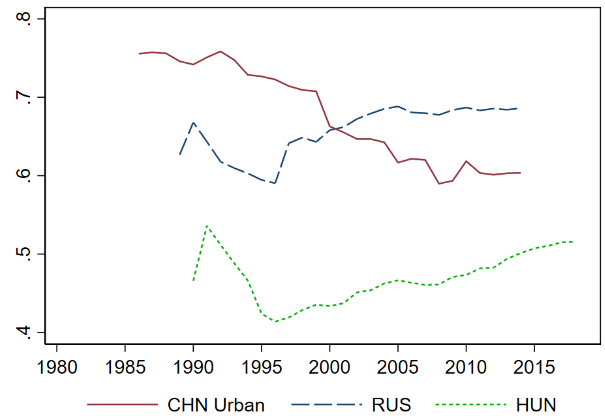
1 Introduction

The increase in female labor supply is an important trend around the world since World War II (Acemoglu, Autor, and Lyle, 2004). Labor supply could be measured by the labor force participation rate (LFP), which is the ratio between the labor force and the working-age population. China, despite having a high female LFP (FLFP), has experienced a significant drop in the FLFP since the 1990s. The decline of FLFP in urban China is significant as it has nosedived nearly 15% from 1990 to 2009 for women above age 25. This pattern is unusual as it is different from typical western countries, former socialist countries, and other eastern Asian countries or regions. In many western countries, FLFP has continued growing or reached its peaks. In many former socialist countries where FLFP was also high before the 1990s, FLFP tended to be negatively shocked by the economic transition but recovered to their original levels after years (Smith, 2011). As for eastern Asian countries or regions, despite the cultural and industrial similarity, there is no similar pattern to China (Figure 1.1).

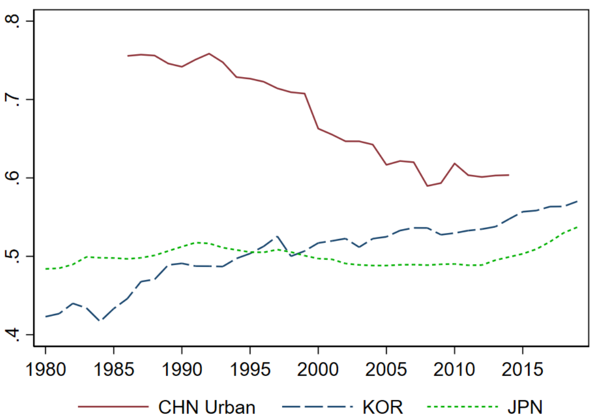
One may argue that China is in a different development stage and therefore the FLFP may increase in the future following the "U" shape trend (Goldin, 1994; Ngai and Olivetti, 2015). However, the FLFP declines in the "U" trend when the society moves from agriculture to manufacturing. In China, FLFP declined in industrialized urban areas. Another argument is that the declined FLFP could be a long-lasting effect of a more moderate economic transition when compared to the "shock therapy" in Russia. There is some sign of FLFP recovery in recent years. However, it is because people are retiring later. I also show that FLFP has declined much more than males' LFP even adjusted for the transition shock. Therefore, the declined FLFP in urban China is intriguing and worth probing to contribute to the long but still growing literature on female labor supply (Section 2).



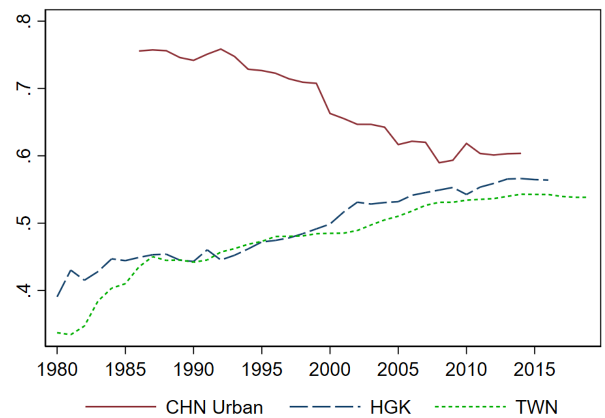
(a) Western Countries



(b) Formal Socialist Countries



(c) Eastern Asian Countries



(d) Eastern Asian Regions

Figure 1.1: Comparison of FLFP (age 25+) between Urban China (CHN), United States (USA), United Kingdom (GBR), Russia (RUS), Hungary (HUN), South Korea (KOR), Japan (JPN), Hong Kong (HOK), and Taiwan (TWN).

Note: The data for China come from Urban Household Survey (UHS). The data for other countries or regions come from International Labor Organization (ILO). ILO data is the national FLFP. They are comparable to the FLFP in urban China as the urbanization rates in these countries or regions are already more than 60-70% since 1980.

I introduce multiple data sets used in this paper and document several important facts in Section 3. I find that the decline of FLFP is driven by women with high school education or lower. With a decomposition exercise, I show that the majority of their declined LFP is due to labor decisions rather than changes in age or education structures. I then explore related economic trends, including changes in market wages and families with various data. I find that despite women's real incomes increasing substantially, the growth rate is slower than that of men's, resulting in a widening gender pay gap across all education levels. A Heckman regression shows that women's participation is negatively affected by their husbands' wages. Within families, the childcare cost has nearly doubled. Moreover, marriage is becoming more assortative.

A life-cycle structural model is developed in Section 4 to study the contributions of each factor among cohorts 1950, 1960, and 1970. In the model, a woman considers her labor decision in a family where the man always works. She chooses to work to increase her consumption, build up her human capital, and pay the childcare cost. On the other hand, she can also choose not to work to enjoy leisure at home and take care of the child by herself to avoid childcare expenses.

In Section 5, I estimate the model with simulated method of moments to uncover preference parameters and wage-related parameters to deal with the selection problem. The model has a good fit, and parameters are verified by comparing elasticities in the model and data, allowing me to study the effect of each channel.

For low-educated women aged 22-49, cohort 1960 and cohort 1970's LFP are on average 4.8% and 8.9% lower than cohort 1950 respectively. The counterfactual study shows that the faster growth of husbands' incomes has a strong negative income effect which pulls down the LFP of the 1960 and 1970 cohorts by 2.6% and 7.3% respectively, explaining a great majority of the falling FLFP. For medium-educated women, the widening gender pay gap is also the main driving force. Further decompositions show that the gender gap in return for experience rather than the gap in wage level is the dominant factor.

The counterfactual study also reveals the important role of family-related channels. A higher childcare cost reduced FLFP by 2% to 4% per year for all women, explaining a significant share of the declined FLFP. An increased assortative marriage makes low-educated women work more but makes medium-educated women work less. A lower fertility rate increases FLFP slightly. Finally, a simple decomposition exercise shows the difference between cohort 1950 and 1970 can explain 36-48% of the decline in FLFP between 1989 and 2009.

This paper contributes to the literature on female labor supply in three aspects. First, structural models have been widely used in studying the female labor supply in developed countries but are less used in developing countries. This paper is among the first few structural models to study China's female labor supply. Second, the classic work of Blau and Kahn (2007) shows a positive relationship between a closing gender pay gap and a higher FLFP. This paper employs China's special gender pay gap trend to study this topic from a novel direction, showing a widening gender pay gap could result in a declining FLFP, extending the horizon of classical models. Finally, this paper highlights that family structure is equally important as income in understanding women's labor decisions. Further discussion is made in Section 6.

2 Literature

FLFP has increased in many countries in the past few decades and has a profound impact on society. Juhn and Potter (2006) and Blau and Kahn (2006, 2013) review the change of FLFP in the U.S. since the 1960s. Olivetti and Petrongolo (2016) review the trend for more industrialized countries.

At the macro level, the change of FLFP could be linked to multiple macro trends such as structural transformation (Ngai and Petrongolo, 2017) and international trade (Yu et al., 2021). At the micro level, the change of FLFP is usually studied within a household setting. As many channels could influence FLFP simultaneously, a structural approach is often necessary. This line of literature dates back to Heckman and Macurdy (1980) and Eckstein and Wolpin (1989). Solid works have been carried out by Attanasio, Low, and Sánchez-Marcos (2008), Eckstein and Lifshitz (2011). They find the rise in education, narrowing of the gender pay gap, and reduced childcare costs are the main driving forces to explain the increase of FLFP in the U.S.

This field is still very active. Blundell et al. (2016) study the effects of policy reforms on FLFP in the U.K. Chiappori, Monica Costa, and Meghir (2018) endogenize the education choice and formation of families. More recently, Albanesi and Prados (2022) have noticed a slowing convergence of LFP since the 1990s among college-educated women in the U.S. They reason that a faster growth of husbands' wages has a negative income effect on FLFP. Coglianese (2018) extends the field by showing that the growing wages of wives explain the high short-term drop of men in the U.S.

The decline of FLFP in China has also attracted much attention but researchers are often hindered by limited data. Feng, Hu, and Moffitt (2017) use the richest micro data known so far to show a reliable trend of labor supply between 1988 and 2009, confirming the existence of a declined FLFP in urban China. To understand this trend, some channels have been studied but most of the existing research focuses on one channel each time. Structural models are even less used with the exception of Jin (2016) and Gao (2020), who focus on near-retirement women and pension policies.

Market wage is one main channel that could influence LFP. Chi and Li (2008, 2014) document that the gender pay gap is widening in China, which is an ever more unusual trend. The relation between a *closing* gender pay gap and rising FLFP is established for some western countries. However, for studies on China's FLFP, the scope is often limited to the women's own incomes and not extended to husbands' incomes (Hare, 2016).

Connelly (1992) points out the importance of childcare costs in females' labor decisions. Due to limited data, this topic is less touched for China. Maurer-Fazio et al. (2011) study this factor without directly measuring the childcare cost. They find women work less with the presence of a young child. Song and Dong (2018) measure the childcare cost with a 2010 survey and find that a higher childcare cost reduces females' labor supply. With rich data, this paper measures childcare costs between 1986 and 2009 directly. More importantly, the effect of childcare costs on FLFP could be compared with other channels in a structural model.

Lastly, marriage is becoming increasingly assortative in China. The degree of assortative marriage has been measured with various indicators (Han, 2010; Feng and Tang, 2019), but not linked to the declining FLFP. This paper bridges these two trends. Besides, a new and robust model-based indicator for assortative marriage, proposed by Chiappori, Costa Dias, and Meghir (2020) is used.

3 Data and Facts

3.1 Data and related background

In this study, I use microdata from the Urban Household Survey (UHS). UHS is a repeated cross-section data collected by the National Bureau of Statistics of China. The data range from 1986 to 2014. Between 1986 and 2009, it covers 16 provinces, 390,000 households, and 1.2 million individuals. Between 2010 and 2014 only 4 provinces' data are available, covering 21,000 households and 0.16 million individuals. As these 4 provinces are more developed, I detail the process to clean and reconcile these two data in Appendix B.1 and B.2.

The official retirement age is 60 for men and 50-55 for women in China (depending on occupations). By 25, most people have finished their education. Therefore I focus on people between the ages of 25-54 for cross-year comparison. Ideally, I should also limit the sample to married women but the marital status was not covered by the UHS until 2002. It is also not easy to identify married women within a household as the cohabitation of relatives is common (e.g. I can not tell if a woman is the male house head's wife or sister). However, by age 29, more than 90% of women born before 1974 have married. Therefore, it is safe to assume the data represents married women mostly. (Figure A.1 in appendix).

To investigate FLFP by education groups, I define low-educated as people with junior high-school degrees or below. Medium-educated is people with senior high-school degrees or equivalent. High-educated are people with some college education or more.

To capture the change in educational levels, I focus on young women aged 25-29. In 1987, nearly two-thirds of young women of these ages are low-educated in urban China. Only 2% of young women at these ages have some college education or more. In the next twenty years, the education of women has been greatly improved. By 2010, more than 30% of women aged 25-29 are high-educated. However, nearly 50% of women at these ages still have junior high-school degrees or below. Low and medium-educated women still make up a large share of urban China even in 2010 (Table 3.1). Therefore, the labor market outcome for these two groups is still substantial and worth studying.

Table 3.1: Share of women in each education group in urban China (age 25-29).

	1987	1995	2000	2005	2010
Low-educated	65%	67%	62%	55%	45%
Medium-educated	33%	24%	25%	24%	21%
High-educated	2%	9%	13%	21%	34%

Note: The data come from the China population censuses. The majority of young women are low or medium-educated in the urban area, even in 2010.

With the UHS data, four relevant facts and their potential influence on FLFP are discussed in the following sections.

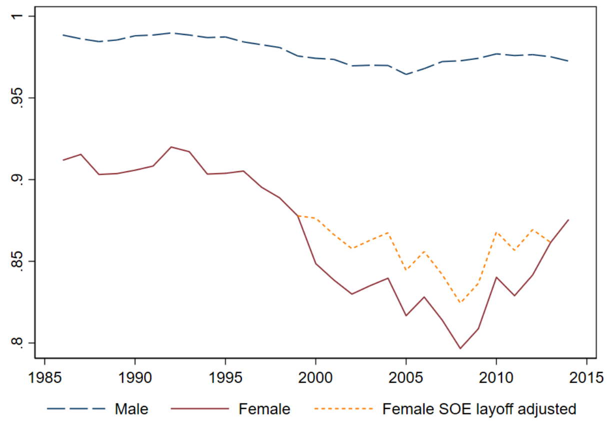
3.2 Fact 1: Changes in FLFP

Figure 3.1a shows trends of LFP by gender across years for people aged between 25 and 54. From 1990 to 2009, men's LFP falls by just 1% while women's LFP dropped by nearly 10%. Since 2009, the FLFP has increased but it is mostly due to delayed retirement. see Appendix B.3 for details.

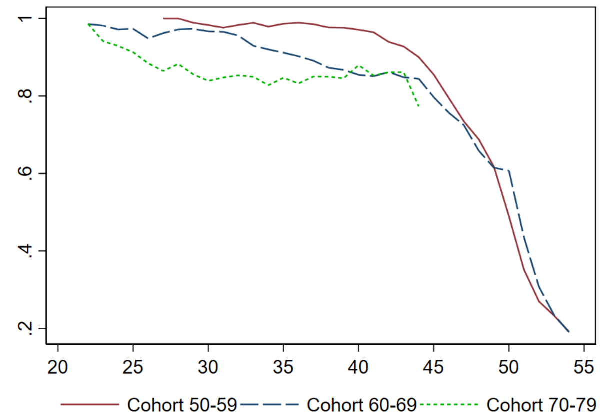
Like many former socialist countries with centrally planned economies, China has experienced a transition period, which is often associated with a decline in LFP. Between 1998 to 2004, there was a massive layoff across state-owned enterprises (SOE) and female workers suffered disproportionately. Many of them chose or were forced to leave the labor market. The orange short dashed line adjusts this shock. Still, the FLFP has dropped nearly 7%. See Appendix B.4 for details of this adjustment.

Figure 3.1b, 3.1c, and 3.1d break the FLFP by cohorts (cohort 1950-1969, cohort 1960-1969, cohort 1970-1979) and education groups. It is clear that the decline of FLFP is driven by low and medium-educated women while the LFP of high-educated women remains high. For the low-educated group, their LFP dropped from nearly 100% to 85% for ages 30-40. For the medium-educated group, their LFP also dropped from nearly 100% to 93% for ages 30-40. However, for both groups, FLFP increased in their 50s, corresponding to the increased LFP after the year 2009 in Figure 3.1a.

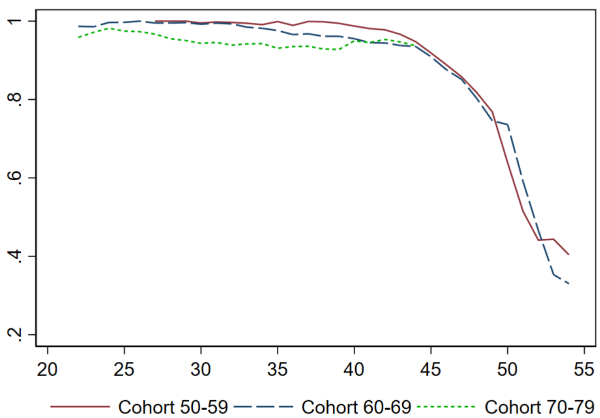
We can also notice year effects in the cohort view. The decline of LFP for cohort 1960-1969 is from age 30-40, roughly corresponding to the year 1995-2005. The decline of LFP for cohort 1970-1979 is from age 20 to 30, also falling between the years 1995 to 2005.



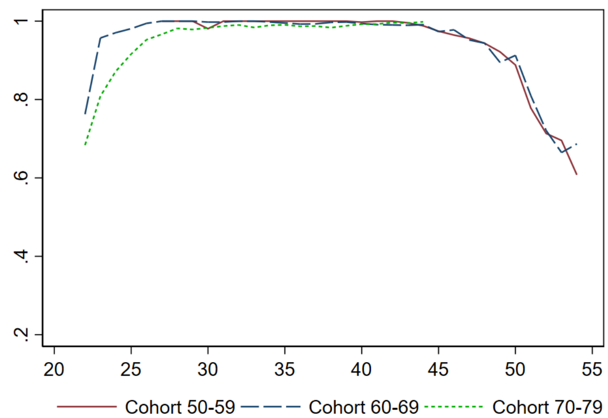
(a) LFP by gender across years



(b) FLFP of low-educated



(c) FLFP of medium-educated



(d) FLFP of high-educated

Figure 3.1: FLFP by years and by cohorts.

Note: FLFP is driven down by low and medium-educated women. FLFP is not further adjusted for the SOE layoff in the cohort view. Despite the large aggregated effect in the year view, the SOE layoff would just increase FLFP by 0.01-0.05% to certain ages.

3.2.1 Decomposition of the FLFP

Graphic evidence suggests that labor decisions of low and medium-educated women play an important role in driving down the FLFP. A caveat is that the drop in FLFP could be due to changes in demographic structures like age or education. For example, the decline of FLFP in the U.S. after 2010 is largely due to an aging society (Krueger, 2017). I decompose the change of FLFP between 1989 and 2009 by Equation (1), where $FLFP_t^i$ is the FLFP of group i in time t. I consider two scenarios: 1) groups are divided by age ($i = age$); 2) groups are divided by both age and education level ($i = age \times education$). W^i is the corresponding population share. The first item in Equation (1) represents changes in the total FLFP explained by changes in the FLFP of certain groups, keeping demographic structures the same.

The first column of Table 3.2 shows that, between 1989 and 2009, FLFP decreased by 9.5%. The second column shows that, if the age structure remained the same, the change in FLFP of each age group would lower total FLFP by 6.4%, explaining 67% of the decline. The fourth column shows that, if the age and education structure remained the same, the change in FLFP of each age and education group would lower total FLFP by 11.3%, meaning that the improvement in education offsets some declines due to labor decisions. Similar patterns could be found if the whole period is separated into two decades.

Combined with graphic evidence, we can confirm that the declining FLFP in urban China is mostly due to the labor choice of low and medium-educated women.

$$\Delta FLFP_{t1}^{t0} = \sum_i \Bigg[\underbrace{\left(FLFP_{t0}^i - FLFP_{t1}^i\right) W_{t0}^i}_{\text{labor decision}} + \underbrace{\left(W_{t0}^i - W_{t1}^i\right) FLFP_{t0}^i}_{\text{demographic structure}} \Bigg] \tag{1}$$

Table 3.2: Decomposition of FLFP 1989-2009 (age 25-54)

Time Period	Δ FLFP	$i = age$		$i = age \times education$	
		labor	demographic	labor	demographic
1989-2009	9.5%	6.4%	3.1%	11.3%	-1.8%
		(67%)	(33%)	(119%)	(-19%)
1989-1999	2.6%	1.4%	1.2%	3.3%	-0.7%
		(54%)	(46%)	(126%)	(-26%)
1999-2009	6.9%	4.7%	2.2%	6.8%	0.1%
		(68%)	(32%)	(99%)	(1%)

Note: Relative explanatory power of labor decision or demographic structure is in the parenthesis.

3.3 Fact 2: Gender Pay Gap

Labor income has grown substantially in China, especially after the economic transition after the year 2000. An increase in women's *own* labor income would encourage them to work more. However, an increase in their husbands' income would act as a non-labor income. An increase in non-labor income would only have an income effect and discourage women to work.

In reality, the growth of men's income is in general faster than women's. Figure 3.2 illustrates the average annual income (the working hour is not asked in most UHS surveys) of low and medium-educated men and women. The annual income has increased fast from the old cohort to the younger cohort for both men and women. However, women's annual incomes are lower than men's both in wages (intercept) and return to experience (slope).

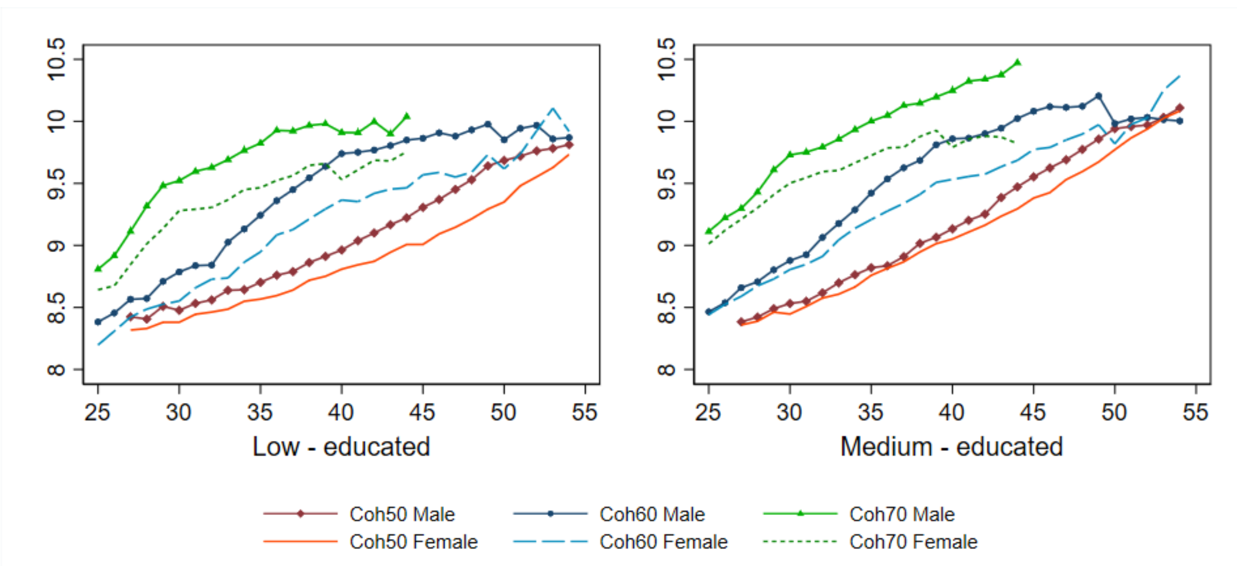


Figure 3.2: Logarithm of real income of low and medium-educated people (age 25-54, in 2009 price).

Figure 3.3 shows the ratio between women's and men's income after controlling for the experience. We can see that the pay gap is widening for all education groups since 1990, especially between 1995 and 2005. This trend is robust when industry and province fixed effects are controlled (Figure A.2 in appendix). Admittedly, the annual income is not the perfect measure for the gender wage gap as the working hours of men and women may be different. However, hours are only surveyed in a few years in the UHS survey.

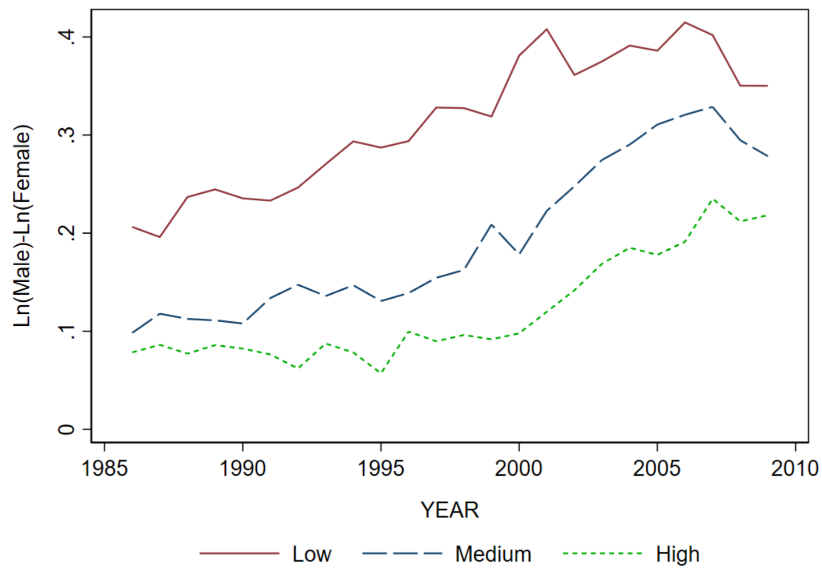


Figure 3.3: Gender pay gap across years by education groups (age 25-54).

Beyond the graphic evidence, the relationship between women's wages and their husbands' wages and FLFP could be explored directly at the household level as Equation (2). Because I only observe wages for working women, a Heckman regression (Equation 3) is used to predict the potential wages for women ($\widehat{\text{Own Wage}}$). It controls for the presence of young children, age, year, city, and education fixed effects. The results are shown in columns (1) and (2) in Table 3.3. With the potential wages, the coefficients of Equation (2) are shown in column (3) and odd ratios are converted to elasticity in column (4) for an easier interpretation: A woman's potential wage is positively related to her participation and her husband's wage is negatively related to her participation.

These elasticities are estimated from the data directly without using identification strategies. To the best of my knowledge, there is no relevant policy shock or instrument variable available in China's setting to achieve clear identification. However, their scale is in line with reduced-form or structural estimation of US/UK like Blau and Kahn (2007) and Blundell et al. (2016).

At a more aggregate level, regressions also show that the city-level gender wage gap is negatively related to local FLFP (Tabel A.1 in the appendix).

$$\text{logit(FLFP=1)} = \beta_1 \widehat{\text{Own Wage}} + \beta_2 \text{Husband's Wage} + \text{Control} + \mu_{\text{age}} + \mu_{\text{year}} + \mu_{\text{city}} + \mu_{\text{edu}} \tag{2}$$

$$\text{Own Wage} = \beta_1 \text{Husband's Wage} + \text{Control} + \mu_{\text{age}} + \mu_{\text{year}} + \mu_{\text{city}} + \mu_{\text{edu}} \tag{3}$$

Table 3.3: Household-level Wage and FLFP

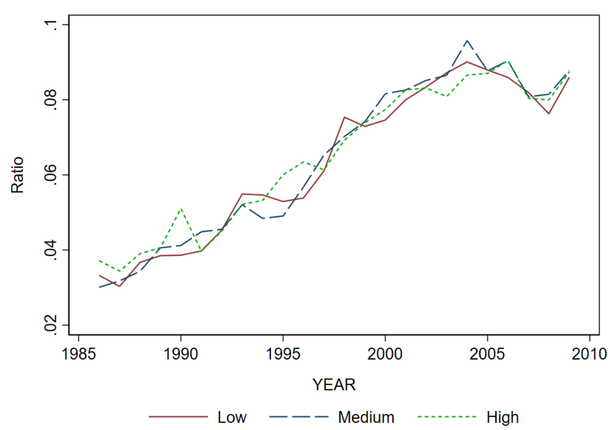
Variables	(1) Selection	(2) Own Wage	(3) FLFP	(4) Elasticity in (3)
$\widehat{\text{Own Wage}}$			1.000***	0.783***
			(0.000)	(0.013)
Husband's Wage	-0.000***	0.348***	1.000***	-0.391***
	(0.000)	(0.000)	(0.000)	(0.007)
Young Child	-0.126***	422***	0.615***	
	(0.000)	(0.000)	(0.000)	
Constant	0.980***	2377***	5.410***	
	(0.000)	(0.000)	(0.000)	
Observations	141,695	141,695	141,695	
(Pseudo) R-squared	0.229	0.541	0.340	
Age FE	yes	yes	yes	
Year FE	yes	yes	yes	
City FE	yes	yes	yes	
Education FE	yes	yes	yes	

Note: Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The sample is limited to women aged 25-54 with a spouse.

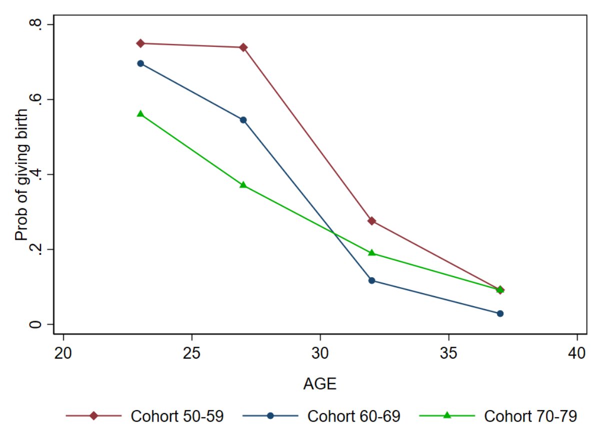
3.4 Fact 3: Higher Childcare Cost and Lower Fertility Rate

Taking care of young children is a duty that often falls on mothers. I define childcare cost as the kindergarten-related fee for children under 6 years old. I find the childcare cost as a ratio to household expenditure has increased from the early 1990s to the early 2000s in Figure 3.4a. Before the 1990s, childcare was widely provided as employee welfare in many SOEs and hardly cost households a penny. During the economic transition in the 1990s, such welfare was no longer universal and many parents turned to commercial kindergartens. A higher childcare cost may make some mothers consider it more cost-efficient to stay at home and take care of the child by themselves.

Meanwhile, fertility rates, measured by the probability of giving birth at a given age, have also waned (Figure 3.4b), which may free women from the family burdens and allow them to work more. When discussing the fertility rate in China, one has to consider the "One-Child Policy", which requires most families to have at most one child. This policy is implemented gradually in the early 1980s and lifted in 2015 so some women in cohort 1950 are affected. Limited data prevents meaningful adjustment. However, as the focus of this paper is on urban women who tend to have a lower fertility rate, the effect of this policy may not be too strong. Besides, there seems no significant cutoff in the fertility rate in the data.



(a) Childcare cost/Household expenditure.



(b) Probability of giving birth.

Figure 3.4: Children-related trends.

Having a child could have more influence on the mother's labor decision other than childcare costs. Taking care of a child is also more than paying kindergarten fees. For example, it may have a "birth penalty" on income. However, as the UHS is not panel data, these effects could not be clearly identified and therefore have to be left out of the model.

3.5 Fact 4: Increased Assortative Marriage

The last fact I looked into is increased assortative marriage. Assortative marriage means people marry others with similar characteristics. This study focuses on the assortative marriage of education. This trend may influence women's labor decisions by affecting husbands' income. For low-educated women, assortative marriage means they are less likely to marry medium or high-educated husbands, who tend to earn more than low-educated men. In that way, their non-labor income of them decreases, which *encourages* them to work more. On the other hand, this trend may allow medium-educated women to *work less* as they are less likely to marry low-educated men.

To measure the degree of assortative marriage, Chiappori, Costa Dias, and Meghir (2020) propose the Separable Extreme Value index (SEV) which is the log value of the ratio between assortative marriage and non-assortative marriage. It can control the change in the population's education and it is based on a transferable utility model, thus more robust than early indicators. A higher value of the SEV index means a higher degree of assortative marriage. With this measurement, I find that assortative marriage has increased by cohort for all education levels (Table 3.4).

Table 3.4: Degree of assortative marriage (SEV index).

	Low-educated	Medium-educated	High-educated
Cohort 50-59	4.19	2.39	3.98
Cohort 60-69	4.25	2.47	4.43
Cohort 70-79	4.70	2.62	4.51

Note: The data come from the 2005 population census. Higher values mean higher degrees of assortative marriage. See Appendix B.5 for details.

4 Model

The sections above discuss three facts and interacting channels which could explain the decreased LFP of low and medium-educated women: widening gender pay gap, higher childcare costs, and more assortative marriage. In this section, I propose a life-cycle model to quantitatively study the contribution of each channel.

As mentioned, the drop in LFP is due to low and medium-educated women. Therefore, I focus on these two groups and assume high-educated women always work. To make full use of the data and make the life-cycle fertility pattern more realistic, I begin the life-cycle from age 22 rather than 25. As shown in Figure 3.1, the LFP of low and medium-educated women are almost equally high between 22 and 25.

4.1 Outline of the Model

A woman of cohort θ enters the economy at age 22 with a given education level $e \in \{L, M\}$. She forms a family at this point with a man of the same age [1] and education level $\tilde{e} \in \{L, M, H\}$ (all variables related to the husband are distinguished by "$\sim$"). The probability for a couple $(e, \tilde{e})$ to form a family, or the degree of assortative marriage, is a cohort-dependent matrix Φ^{θ}. Once a family is formed, it would not dissolve.

The woman's age is denoted by a and she is going to maximize her own lifetime utility by making labor decision p. She knows her income and decides whether to work or not. She retires at 55 and left the model. This maximum age is denoted by $\bar{a}$.

At each period, a childless woman has probability Ψ_a^{θ} to have one and at most one child. The child needs to be taken care of for 6 years. Whether the woman has had a child at a certain age is denoted by N_a and whether the child is young is denoted by χ. The family could pay $\kappa_{\theta+a}$ (κ depends on the calendar year, which equals cohort year plus the women's age) for market childcare service or the mother needs to stay at home to look after the child. The timeline of the model is illustrated in Figure 4.1.

4.2 Preference and Budget Constraint

The preference of a woman is Equation (4), following Attanasio, Low, and Sánchez-Marcos (2008).

$$u(C_a, p_a; n_a, e) = \frac{(C_a/n_a)^{1-\rho}}{1-\rho}\exp[(1-p_a)(\gamma_1^e)] - p_a\gamma_2^e \tag{4}$$

[1] The average age difference between a husband and a wife in UHS data is 2.5. The median age difference is 2. Before 2003, more than 90% of all women have married by the age 25-29 (Figure A.1).

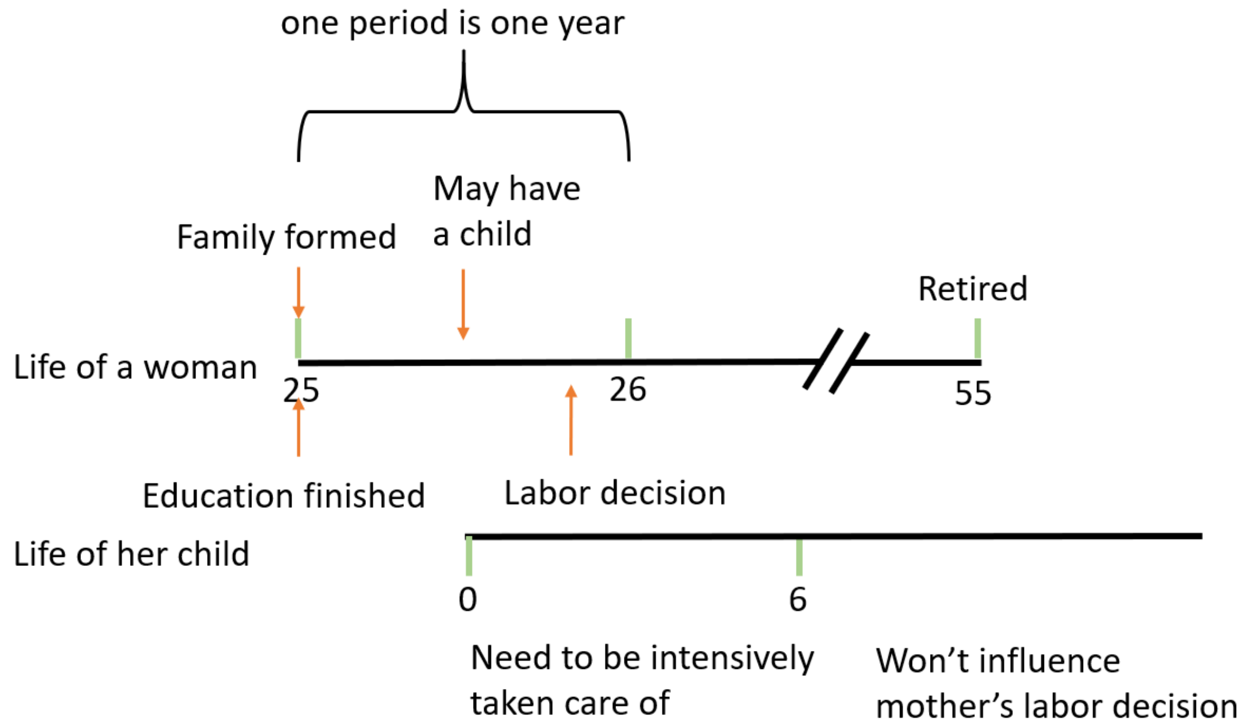


Figure 4.1: Timeline of the model

C_a is the household consumption. To determine the women's consumption share, I use the"OECD-modified" equivalence scale (n_a), which is an empirical constant to determine an individual's consumption as the share of the household consumption. n_a is 1.5 for a family of two adults and 1.8 for a family of two adults and one child[2]. $p_a \in \{0, 1\}$ is her labor decision. ρ is the risk aversion coefficient. γ_1 and γ_2 measure her leisure from not working.

The woman makes decision of p_a based on state variable $Z_a(S_a, y_a, \tilde{y}_a, N_a, \chi)$ at each period to maximize her lifetime utility:

$$V_a^{\theta}(Z_a) = \max_{\{p_\tau\}_{\tau=a,\ldots,\bar{a}}} E\left\{ \sum_{\tau=a}^{\bar{a}} \beta^{\tau-a} u(C_a, p_a; n_a, e) | Z_a \right\} \tag{5}$$

under the budget constraint:

$$C_a + p_a \kappa_{\theta+a} \times \mathbb{1}(\chi = 1) = y_a p_a + \tilde{y}_a \tag{6}$$

[2]See OECD website: http://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf

y_a and $\tilde{y}_a$ are the annual income of women and men. By letting consumption equal to income, I assume there is no saving and borrowing as in (Eckstein and Lifshitz, 2011; Eckstein, Keane, and Lifshitz, 2019). I plan to include saving in the future.

4.2.1 Income Process

The annual income of a woman is determined by a Mincer-type equation:

$$\ln y_a = b_0 + b_1 S_a + b_2 S_a^2 + \epsilon_a \tag{7}$$

$$\begin{aligned} S_a =& S_{a-1} + p_{a-1}, \quad \text{if } p_{a-1} = 1 \\ S_a =& (1-\delta) S_{a-1}, \quad \text{if } p_{a-1} = 0 \end{aligned} \tag{8}$$

y is the observed annual income of the woman. S_a is her accumulated working experience at age a. I assume everyone begins to accumulate experience only after age 22 as young people are often loosely attached to certain jobs or occupations. If the woman works, she accumulates one unit of experience. Otherwise, her experience depreciates by δ.

As y is the observed women's income, the estimation of b could be biased due to selection. This problem could be solved either outside the model using a Heckman model (Chiappori, Monica Costa, and Meghir, 2018) or inside the model with structural estimation (Eckstein and Lifshitz, 2011; Blundell et al., 2016). Here I follow the second approach by estimating the variance of an i.i.d. error term or the variance of the ability to draw $\epsilon_a \sim N(0, \sigma)$. [3]

The husband's income is also determined by a Mince-type equation. I assume the husband is always working (therefore experience is just age) and there is no variance to simplify the problem:

$$\ln \tilde{y}_a = \tilde{b}_0 + \tilde{b}_1 a + \tilde{b}_2 a^2 \tag{9}$$

[3] Keane and Wolpin (2009) discussed a similar environment where a mother faces a trade-off between working or staying at home to take care of the children. She goes to work if her draw of ϵ is higher than her reserved wage draw ϵ^*.

5 Estimation, Results and Counterfactual Studies

5.1 Estimation and Results

The model is used to study the changes in FLFP among three cohorts: cohort 1950-1959 (cohort 50), cohort 1960-1969 (cohort 60), and cohort 1970-1979 (cohort 70).

ρ and β are set as 1.5 and 0.98 as in the literature. Parameters in husbands' income function ($\tilde{b}$) are estimated from the data directly (Table A.2 in the appendix). Φ (degree of assortative marriage), Ψ (fertility rate), and κ (childcare cost) are calculated from the UHS data or the census data.

The rest parameters are estimated with the simulated method of moments (SMM). Women's income parameters ($b_{0,1,2}, \delta, \sigma$) are estimated for each cohort and education group separately. Preference parameters (γ_1, γ_2) are estimated *jointly* across cohorts for each education group so that changes in FLFP are not due to differences in preferences.

The target moments are LFP and annual incomes. There are 56, 66, and 46 moments for three cohorts respectively and the model is over-identified. The estimator $\hat{\Theta} = (b_{0,1,2}, \delta, \sigma, \gamma_1, \gamma_2)$ is defined by:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \left\{ \sum_{k}^{K} [(M_k^d - M_k^s(\Theta))^2 / \mathrm{Var}(M_k^d)] \right\} \tag{10}$$

where k is the kth moment for simulation. M^d is the moment in data and $M^s(\Theta)$ is the simulated moment with parameter Θ. The weighting matrix is the inverse matrix of the variance of data moments.

The estimation results for low and medium-educated are listed in Table 5.1 and 5.2 respectively. χ^2 is the Pearson cumulative test statistic and it is smaller than the corresponding critical value in all cases. Therefore models are not rejected in overidentification tests.

In general, women of the young cohort have higher wages, higher returns of experience, and lower depreciation rates than the old cohort. The fit of the model is shown in Figure 5.1. The fit for low-educated women is very good while the fit for medium-educated women is less precise, partially due to the change in FLFP being smaller.

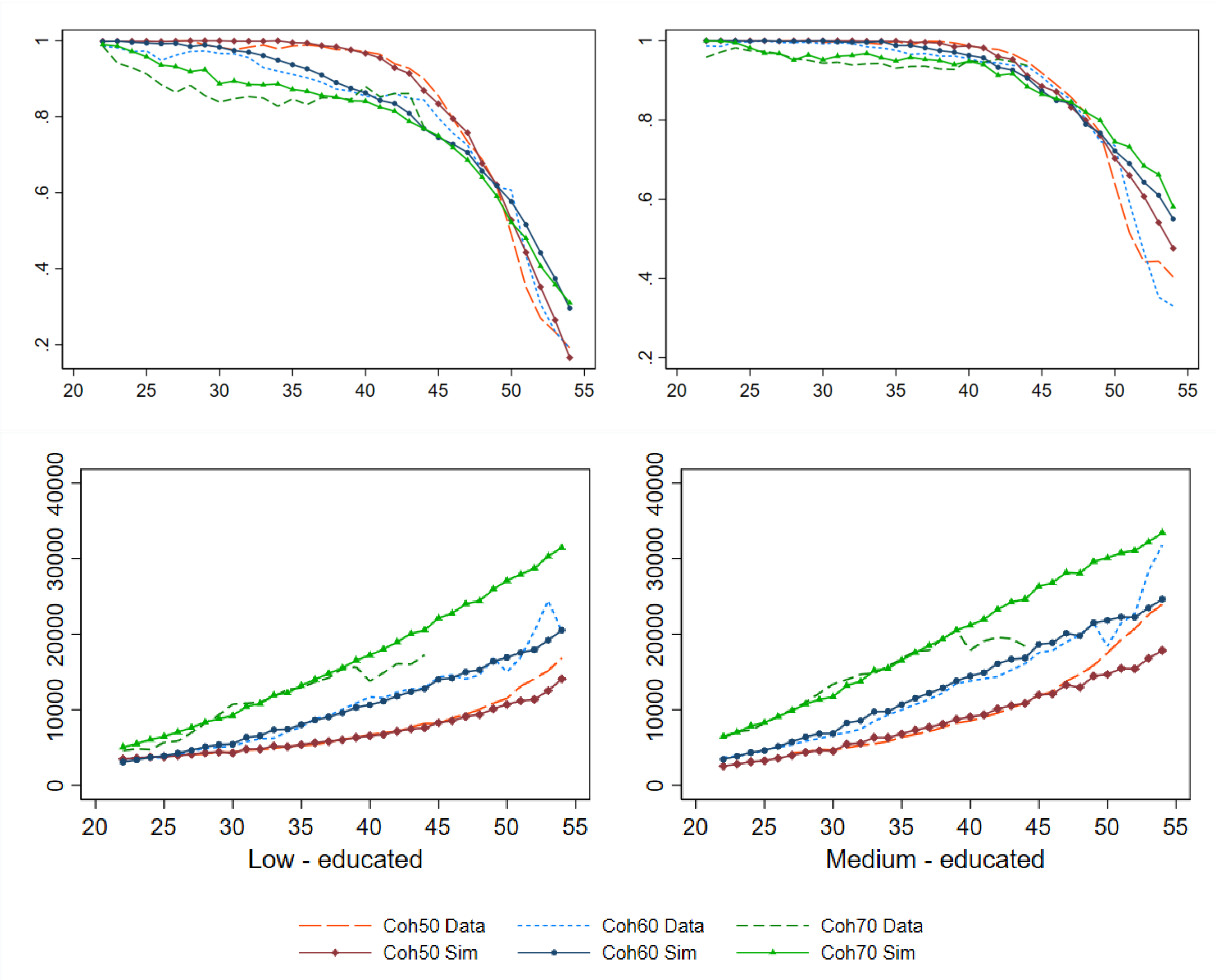


Figure 5.1: Fit of low and medium-educated women's LFP and income (in 2009 price).

5.2 Model Validity

The parameters are verified by comparing implied extensive labor elasticities from the model with elasticities estimated from the data. For elasticities in the model, I calculate the response of FLFP to the change in wage levels, which is known as "life-cycle Marshallian elasticity" in the literature (Attanasio et al., 2018). One example is shown in Table A.3 in the appendix. Elasticities in the data are estimated as Equation (2) and (3) but for dif-

Table 5.1: Estimation results of low-educated women.

	Cohort 50	Cohort 60	Cohort 70
b_0	8.031	7.899	8.455
	(0.000)	(0.001)	(0.006)
b_1	0.033	0.087	0.091
	(0.000)	(0.000)	(0.000)
b_2	0.000	-0.001	-0.001
	(0.000)	(0.000)	(0.000)
δ	0.167	0.150	0.019
	(0.025)	(0.022)	(0.033)
σ	0.471	0.492	0.317
	(0.007)	(0.009)	(0.016)
γ_1		-0.186	
		(0.002)	
γ_2		0.000	
		(0.000)	
χ^2	17.4	30.9	20.6
$\chi^2_{critical}$	66.3	77.9	54.6

Table 5.2: Estimation results of medium-educated women.

	Cohort 50	Cohort 60	Cohort 70
b_0	7.624	7.946	8.675
	(0.000)	(0.001)	(0.007)
b_1	0.094	0.111	0.094
	(0.000)	(0.000)	(0.001)
b_2	-0.001	-0.002	-0.001
	(0.000)	(0.000)	(0.000)
δ	0.040	0.077	0.000
	(0.006)	(0.012)	(0.012)
σ	0.639	0.608	0.387
	(0.005)	(0.007)	(0.012)
γ_1		-0.178	
		(0.003)	
γ_2		-0.000	
		(0.000)	
χ^2	50.6	45.5	13.2
$\chi^2_{critical}$	66.3	77.9	54.6

Note: Standard error in brackets. b_0, b_1, b_2: coefficients in the income equation. δ: experience depreciation rate. σ: variance of ability draw. γ_1, γ_2: preference parameters (estimated jointly across three cohorts). χ^2 is the Pearson cumulative test statistic and χ^2_{critic} is the corresponding critical value at 95% confident level.

ferent age and education groups separately. Again, these results are not identified from policy shocks or with instrument variables but their scale is in line with the literature.

The first line in Panel A of Table 5.3 compares average elasticities in the model and data for low-educated women between ages 25-54. The average elasticity to income (own elasticity) implied by the model is 0.47 which falls into the estimation interval from the data ($[0.13, 1.08]$). The average elasticity to husbands' income (cross elasticity) implied by the model is -0.21 and it also falls into the estimation interval from the data ($[-0.54, -0.13]$). Three points are worth noting: 1) In general, most point estimators from the model fall into interval estimations from the data. 2) All interval estimations of ages 45-54 are not significant from 0 therefore they cannot compare with point estimators from the model (marked with "N/A"). 3) Interval estimations of cross elasticity are narrower than the own elasticities. Similar patterns are also found for medium-educated women in Panel B.

Table 5.3: Elasticity in the model and data

	Own Elasticity			Cross Elasticity		
Age group	Model	Data 95% C.I.	Within	Model	Data 95% C.I.	Within
Panel A: low-educated						
25-54	0.47	[0.13, 1.08]	yes	-0.21	[-0.54, -0.13]	yes
25-34	0.21	[0.24, 1.33]		-0.07	[-0.73, -0.17]	
35-44	0.47	[0.16, 0.70]	yes	-0.18	[-0.38, -0.12]	yes
45-54	0.74	[-1.18, 1.42]	N/A	-0.39	[-0.68, 0.42]	N/A
Panel B: medium-educated						
25-54	0.14	[0.04, 0.16]	yes	-0.10	[-0.10, -0.04]	yes
25-34	0.09	[0.06, 0.19]	yes	-0.07	[-0.13, -0.05]	yes
35-44	0.16	[0.11, 0.22]	yes	-0.10	[-0.14, -0.08]	yes
45-54	0.16	[-0.33, 0.06]	N/A	-0.12	[-0.05, 0.14]	N/A

Note: "Within" is marked with "yes" if the point estimator of the model falls into the interval estimation (95% confidence interval/C.I.) from data. "Within" is marked with "N/A" if the interval estimation includes 0.

The parameters are also verified with non-targeted moments of high-educated women. Their LFP is used as non-targeted moments. Instead of estimating parameters structurally by targeting these moments, I combine non-structurally estimated parameters and already estimated parameters to generate their LFP and to check the goodness of fit. I estimate their income parameters with Equation (9) as the selection problem is not a major concern for high-educated women. The remaining parameters, namely income variance, depreciation rate, and utility parameters are assigned to medium-educated groups. The combination of these parameters generates a good fit of LFP for high-educated groups before age 50 (Figure A.3 in the appendix).

5.3 Counterfactual Studies

With the model, I could now study the contribution of each channel. I denote the LFP of cohort 50-59 as L_{50}, LFP of cohort 60-69 as L_{60}, and LFP of cohort 70-79 as L_{70}. I then substitute parameters of cohort 50-59 with parameters of cohort 60-69 to find the counterfactual LFP L'_{50} (what would the LFP of cohort 50-59 be if they face the same market condition of cohort 60-69?). Similarly, I give cohort 50-59 parameters of cohort 70-79 to find the counterfactual LFP L''_{50}.

As there are multiple channels, there are two approaches to studying their effects. One is to study each channel one by one. The other is to study each channel one *on the top* of each one. Because the utility function is non-homothetic, people prefer leisure more when they have more income and consumption. Therefore, the effects of family-related channels are subject to the presence of income channels. For example, if low-educated women in cohort 50 only face the same childcare cost as cohort 60, they would reduce labor supply by 0.1% on average. However, if they can earn as much as cohort 60, they would reduce labor supply by 2.4% facing the same increase in childcare costs.

Because the change in income is the first-order effect, I would study the effects of family-related channels upon income channels. One potential problem is that the order of putting channels would influence their effects. In the current specification, assortative marriage, fertility rate, and childcare cost are added in sequence to reflect their natural order. However, altering their orders would not change the sign and scale of each channel significantly as long as these family-related channels are added after income channels (Table A.4).

Figure 5.2 shows LFP changes in the life-cycle response to several channels for low-educated women between cohort 1950 and 1960. Graphs on the left column display the counterfactual LFP in dashed lines and graphs on the right column compare $L_{70} - L_{50}$ and $L''_{50} - L_{50}$, highlighting the explaining power of each channel.

On average, LFP is 8.9% per year lower than cohort 1950 between 22-49 for cohort 1970. The first row of graphs illustrates the effect of changes in couples' income. We can see that the unbalanced growth rate between men and women, or the gender gap in return of experience, has a huge negative effect on women's LFP (the green dashed line), which on average reduces FLFP by 17.1%. Women's higher wages (gap in wage) have offset this trend partially, increasing FLFP by 9.8% on average. The net effect of changes in couple's incomes (return and wage) has lower FLFP by 7.3% each year on average between 22-49, explaining 82% of the declined LFP [4].

The second row of graphs compares the income channel and the family channel. The green dashed line is changes in couples' incomes plus unobservable depreciation rate and variance in ability. Accounting for these two additional channels reduces the overall effect of the income channel by 10%, leaving 30% of the declined LFP to be explained by the family channel.

Within the family channel, changes in assortative marriage and fertility rate actually increase FLFP by nearly 2% per year: increased assortative marriage reduces the probability of low-educated women finding high-educated partners and therefore reducing their non-labor income, encouraging them to work more. Reduced fertility rates also allow them to work more. However, the main driving force in the family channel is the increased childcare cost, which reduces FLFP by 4.4%, explaining nearly 50% of the reduced FLFP on top of the income channel.

[4] Measured by $\frac{7.3\%}{8.9\%} = 82\%$, similar for other calculations.

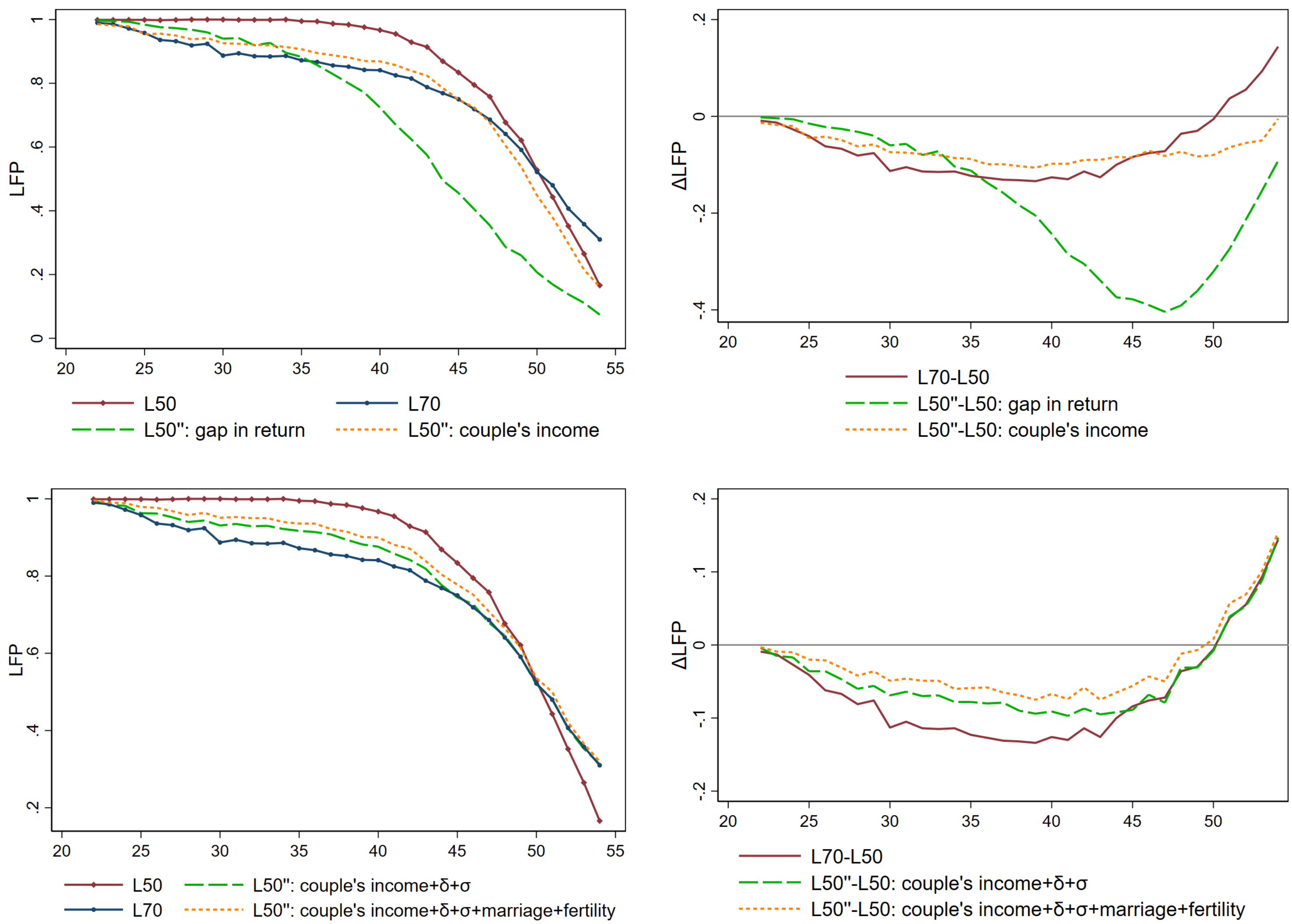


Figure 5.2: Counterfactual LFP for low-educated women between cohort 1950 and 1970.

Left: comparison between levels. Right: comparison between differences. The solid lines are reference LFP and the dashed lines are counterfactual LFP.

Figure 5.3 shows similar patterns for medium-educated women between cohort 1950 and 1970: unbalanced wage growth rate has an even larger negative effect for medium-educated women. Changes in income explain 70% of the declined LFP.

Reduced fertility rates also increase FLFP but the scale is less significant. However, for medium-educated women, increased assortative marriage benefits them, increasing their non-labor income which reduces their labor supply by 0.4% (15% of their reduced LFP). Increased childcare cost also matters for them, which explain almost all of the declined

LFP on top of the income channel.

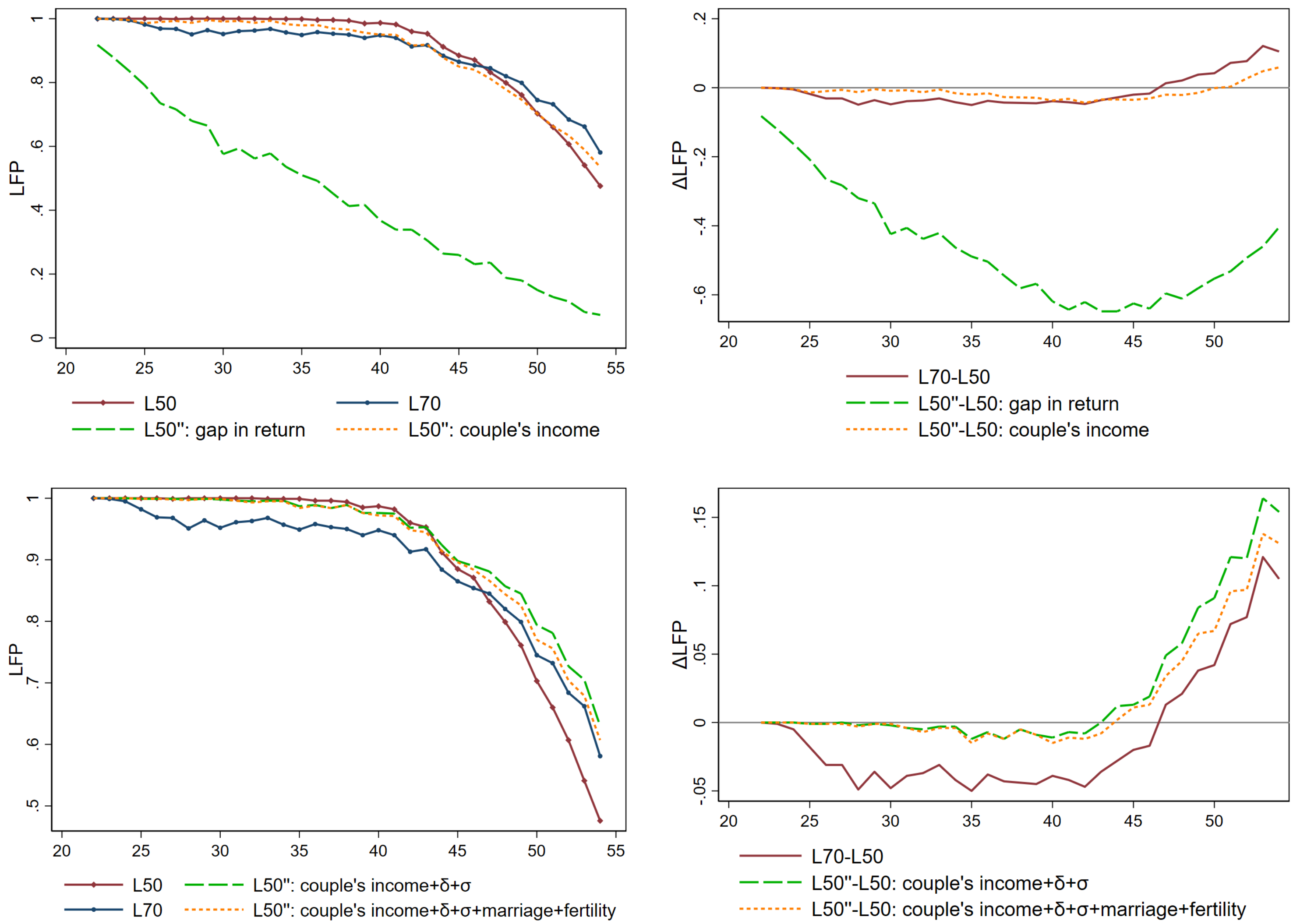


Figure 5.3: Counterfactual LFP for medium-educated women between cohort 1950 and 1970.

Left: comparison between levels. Right: comparison between differences. The solid lines are reference LFP and the dashed lines are counterfactual LFP.

Table 5.4 summarizes quantitative results. The first row of the table shows average changes of LFP between 22-49 in data. I focus on this period because it is when most LFP decline happens and the main focus of this paper.

Three panels show *accumulated* effects of each channel. For example, for low-educated women in cohort 50, I let them have the same gap in return as cohort 60, and their LFP increases by 1.9% on average. I then let them have the same gap in wage *upon* the same gap in return and their LFP decreases by 4.5%. The net effect of changes in couples' incomes reduces FLFP by 2.6%. Similarly, I put in depreciation rate and variance in ability, assortative marriage, fertility rate, and childcare cost gradually to identify the effect of each channel.

Panel A shows the effect of changes in couples' incomes. Except for medium-educated women between cohort 50 and 60, changes in couples' income can explain 50%-80% the declined LFP. Except for low-educated women between cohorts 50 and 60, changes in the gap in return have generated strong negative income effects, suggesting that the unbalanced growth rate between men and women is the main driving force.

Panel B shows the effect of changes in couples' incomes plus unobservable parameters (depreciation rate and variance in ability). For low-educated women, results are close to panel A, suggesting the results are robust. For medium-educated women, results have different signs with panel A, which is worth more investigating.

Panel C shows the effect of changes in family structures. Increased assortative marriages make low-educated women work more but make medium-educated women work less. Lower fertility rates increase LFP for all women. Increased childcare costs reduce LFP for all women except medium-educated women between cohort 50 and 60 and they are the main driving forces in family-related channels.

Table 5.4: Effect of each channel on LFP per year between 22-49.

	Low-educated		Medium-educated	
	Cohort 50-60	Cohort 50-70	Cohort 50-60	Cohort 50-70
Total Change	-4.8%	-8.9%	-0.8%	-2.7%
Panel A:				
Couple's Income	-2.6%	-7.3%	1.0%	-1.9%
gap in return	1.9%	-17.1%	-3.2%	-45.9%
gap in wage	-4.5%	9.8%	4.2%	44.0%
Panel B:				
Couple's Income $+\delta+\sigma$	-2.7%	-6.4%	-0.7%	0.5%
Panel C:				
Family Structures	-2.1%	-2.5%	-0.1%	-3.3%
assortative marriage	0.1%	1.1%	-0.3%	-0.4%
fertility rate	0.3%	0.8%	0.0%	0.1%
childcare cost	-2.4%	-4.4%	0.2%	-2.8%

Note: This table shows the accumulated effect of each channel. Couples' income is the combination of gaps in return to experience (b_{12} and $\tilde{b_{12}}$) and gaps in wage (b_0 and $\tilde{b_0}$). δ: depreciation rate. σ: variance in ability. Family structures combine assortative marriage, fertility rate, and childcare cost.

6 Conclusion

The declining FLFP in China is a strong and unique trend. Though many channels have been studied separately, this paper employs a structural model to study their relative importance together within a uniform framework. The model shows the unique pattern of China's female labor supply could be explained by a standard labor supply model without involving complicated institutional differences between China and other countries.

The counterfactual study shows that faster growth in husbands' incomes has a strong income effect, pulling down the FLFP by 2.6% and 7.3% each year at a young age for low-educated women in cohort 1960 and 1970 respectively, explaining a large share of the falling FLFP. This trend also has explanatory power for medium-educated women, revealing the widening gender pay gap is the main mechanism behind the falling FLFP. Family-related trends have significant but heterogeneous effects on FLFP.

These cohort-level quantitative results could be mapped to year-level FLFP decline following the decomposition of Equation (1). Allowing low and medium-educated women in cohort 1950 in 1989 to have the same FLFP as cohort 1970 in 2009 reduces total FLFP in 1989 by 5%, explaining 48% of the total change between 1989 and 2009 (the total decline in FLFP is 9.5%). From Table 5.4, 75% of these cohort-level changes are explained by the income channel which corresponds to 36% of the total decline in FLFP between 1989 and 2009.

The next question would be why the gender wage gap has been widening? It turns out to be a more complicated question that is beyond the scope of this paper. Some preliminary analysis shows that the gender wage gaps are widening in both manufacturing and service sectors, which is against the structural change theory that women benefit from the structural change as they have comparative advantages in the service sector. Changes in working hours, time allocation, and probability of promotion could be promising approaches but may rely on exploring broader data.

References

ACEMOGLU, D., D. H. AUTOR, AND D. LYLE (2004): "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury," *Journal of Political Economy*, 112, 497–551.

ALBANESI, S. AND M. J. PRADOS (2022): "Slowing Women's Labor Force Participation: The Role of Income Inequality," Working Paper 29675, National Bureau of Economic Research.

ATTANASIO, O., P. LEVELL, H. LOW, AND V. SÁNCHEZ-MARCOS (2018): "Aggregating Elasticities: Intensive and Extensive Margins of Women's Labor Supply," *Econometrica*, 86, 2049–2082.

ATTANASIO, O., H. LOW, AND V. SÁNCHEZ-MARCOS (2008): "Explaining Changes in Female Labor Supply in a Life-Cycle Model," *American Economic Review*, 98, 1517–52.

BLAU, F. AND L. KAHN (2007): "Changes in the Labor Supply Behavior of Married Women: 1980–2000," *Journal of Labor Economics*, 25, 393–438.

BLAU, F. D. AND L. M. KAHN (2006): "The U.S. Gender Pay Gap in the 1990s: Slowing Convergence," *Industrial and Labor Relations Review*, 60, 45–66.

——— (2013): "Female Labor Supply: Why Is the United States Falling Behind?" *American Economic Review*, 103, 251–56.

BLUNDELL, R., M. COSTA DIAS, C. MEGHIR, AND J. SHAW (2016): "Female Labor Supply, Human Capital, and Welfare Reform," *Econometrica*, 84, 1705–1753.

CHI, W. AND B. LI (2008): "Glass ceiling or sticky floor? Examining the gender earnings differential across the earnings distribution in urban China, 1987–2004," *Journal of Comparative Economics*, 36, 243 – 263.

——— (2014): "Trends in China's gender employment and pay gap: Estimating gender pay gaps with employment selection," *Journal of Comparative Economics*, 42, 708 – 725.

CHIAPPORI, P.-A., M. COSTA DIAS, AND C. MEGHIR (2020): "Changes in Assortative Matching: Theory and Evidence for the US," Working Paper 26932, National Bureau of Economic Research.

CHIAPPORI, P.-A., D. MONICA COSTA, AND C. MEGHIR (2018): "The Marriage Market, Labor Supply, and Education Choice," *Journal of Political Economy*, 126, S26–S72.

COGLIANESE, J. (2018): "The rise of in-and-outs: Declining labor force participation of prime age men," *Harvard University. https://scholar. harvard. edu/files/coglianese/files/coglianese_2017_in-and-outs. pdf*.

CONNELLY, R. (1992): "The Effect of Child Care Costs on Married Women's Labor Force Participation," *The Review of Economics and Statistics*, 74, 83–90.

ECKSTEIN, Z., M. KEANE, AND O. LIFSHITZ (2019): "Career and Family Decisions: Cohorts Born 1935–1975," *Econometrica*, 87, 217–253.

ECKSTEIN, Z. AND O. LIFSHITZ (2011): "Dynamic Female Labor Supply," *Econometrica*, 79, 1675–1726.

ECKSTEIN, Z. AND K. I. WOLPIN (1989): "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience," *The Review of Economic Studies*, 56, 375–390.

FENG, S., Y. HU, AND R. MOFFITT (2017): "Long run trends in unemployment and labor force participation in urban China," *Journal of Comparative Economics*, 45, 304 – 324.

FENG, S. AND G. TANG (2019): "Accounting for Urban China's Rising Income Inequality: The Roles of Labor Market, Human Capital, and Marriage Market Factors," *Economic Inquiry*, 57, 997–1015.

GAO, H. (2020): "Social Security and Female Labor Supply in China," .

GOLDIN, C. (1994): "The U-shaped female labor force function in economic development and economic history," .

HAN, H. (2010): "Trends in educational assortative marriage in China from 1970 to 2000," *Demographic Research*, 22, 733–770.

HARE, D. (2016): "What accounts for the decline in labor force participation among married women in urban China, 1991–2011?" *China Economic Review*, 38, 251–266.

HECKMAN, J. AND T. E. MACURDY (1980): "A Life Cycle Model of Female Labour Supply," *Review of Economic Studies*, 47, 47–74.

JIN, W. M. (2016): "Explaining low employment rates among older women in urban China," IFS Working Papers W16/24, Institute for Fiscal Studies.

JUHN, C. AND S. POTTER (2006): "Changes in Labor Force Participation in the United States," *The Journal of Economic Perspectives*, 20, 27–46.

KEANE, M. P. AND K. I. WOLPIN (2009): "Empirical applications of discrete choice dynamic programming models," *Review of Economic Dynamics*, 12, 1 – 22.

KRUEGER, A. (2017): "Where Have All the Workers Gone? An Inquiry into the Decline of the U.S. Labor Force Participation Rate," *Brookings Papers on Economic Activity*, 48, 1–87.

MAURER-FAZIO, M., R. CONNELLY, L. CHEN, AND L. TANG (2011): "Childcare, eldercare, and labor force participation of married women in urban China, 1982–2000," *Journal of Human Resources*, 46, 261–294.

NGAI, L. R. AND C. OLIVETTI (2015): "Structural transformation and the U-shaped female labor supply," 1501.

NGAI, L. R. AND B. PETRONGOLO (2017): "Gender Gaps and the Rise of the Service Economy," *American Economic Journal: Macroeconomics*, 9, 1–44.

OLIVETTI, C. AND B. PETRONGOLO (2016): "The Evolution of Gender Gaps in Industrialized Countries," *Annual Review of Economics*, 8, 405–434.

SMITH, K. (2011): "Labor force participation in the Soviet and post-Soviet Baltic States," *Economic Change and Restructuring*, 44, 335–355.

SONG, Y. AND X. DONG (2018): "Childcare Costs and Migrant and Local Mothers' Labor Force Participation in Urban China," *Feminist Economics*, 24, 122–146.

TIAN, X., J. GONG, AND Z. ZHAI (2022): "The effect of job displacement on labor market outcomes: Evidence from the Chinese state-owned enterprise reform," *China Economic Review*, 72, 101743.

YU, Z., X. WU, M. LI, AND R. GUO (2021): "Import competition and the gender gap in labor force participation: Evidence from China," *China Economic Review*, 69, 101689.

A Supplementary Figures and Tables

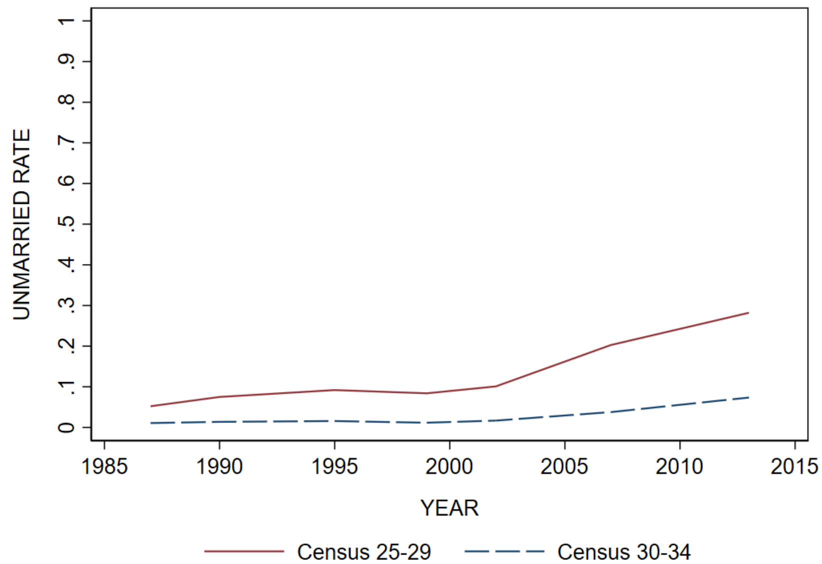
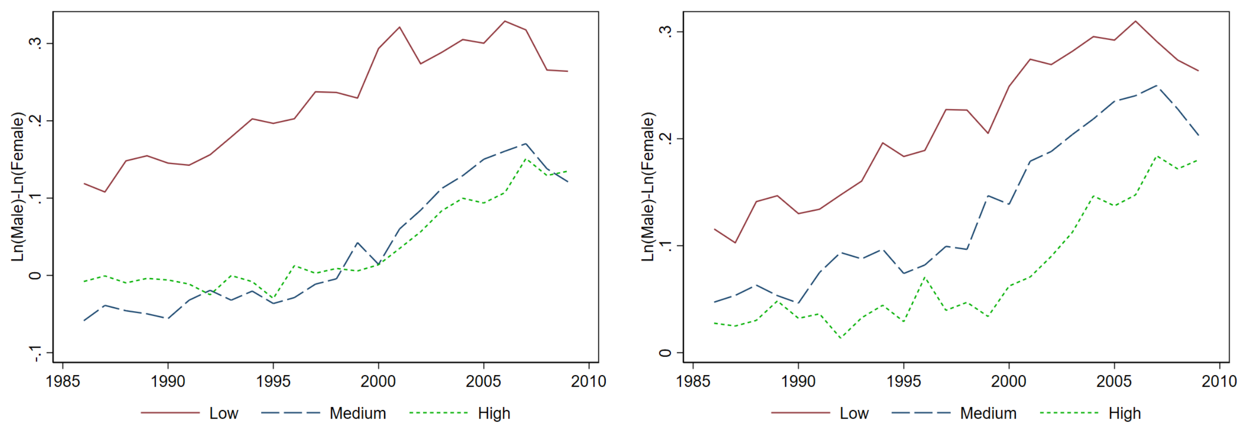


Figure A.1: Never married rate in urban China.

Source: *China Population & Employment Statistics Yearbook*. Between 1995 and 1999, national census data are used instead of urban data, which are not available. (Refer to the main text section 3.1)



(a) Residual wage approach.

(b) Gender dummy approach.

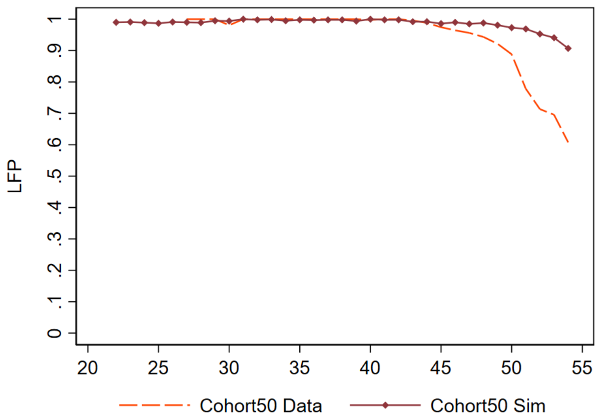
Figure A.2: Gender pay gap in alternative measurements.

Note: (a) measures the pay gap by the difference between residual wages, which are the constant terms in Mincer regressions run for men and women separately. Experience and square of experience are controlled (the latter one is not controlled in Figure 3.3). (b) measures the pay gap by introducing a gender dummy variable in Mincer regressions. Besides experience and square of experience, industries and provinces are also controlled. Referred in main text section 3.3)

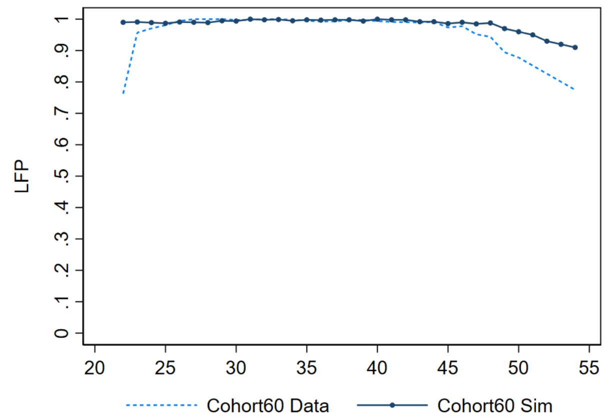
Table A.1: City-level Gender Wage Gap and FLFP

Variables	(1) ln(FLFP)	(2) ln(FLFP)	(3) ln(FLFP)
ln(Male/Female Wage)	-0.273***	-0.220***	-0.029**
	(0.015)	(0.017)	(0.015)
Constant	-0.116***	-0.125***	-0.161***
	(0.003)	(0.003)	(0.003)
Observations	2,949	2,936	2,936
R-squared	0.130	0.466	0.654
City FE		yes	yes
Year FE			yes
Population Weight	yes	yes	yes

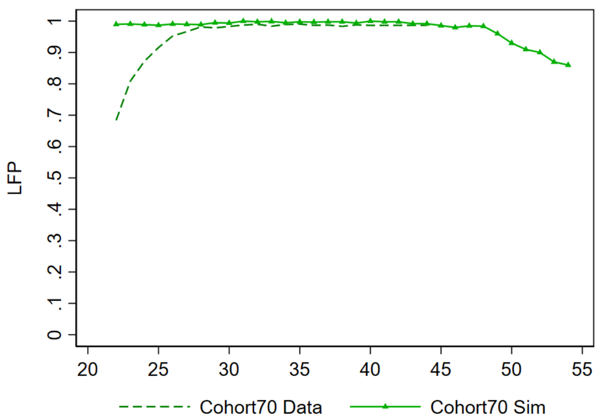
Note: Standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The gender wage gap and the FLFP are negatively related at the city level. The second column is a preferred specification: 1% increase in the gender pay gap decreases FLFP by 0.2% which captures the scale in Figure 3.3. Controlling for the year-fixed effect essentially captures the difference between cities within the same year which explains the much smaller coefficient in the third column.



(a) Cohort 50



(b) Cohort 60



(c) Cohort 70

Figure A.3: Fit of high-educated women's LFP (non-targeted moment).

Note: Income parameters are estimated from Equation (9). Other parameters come from medium-educated women of the same cohort. (Refer to the main text section 5.2)

Table A.2: Estimation of men's earning

		Cohort 50	Cohort 60	Cohort 70
	$\tilde{b_0}$	8.2014	8.0750	8.4155
L	$\tilde{b_1}$	0.0312	0.1068	0.1433
	$\tilde{b_2}$	0.0007	-0.0013	-0.0016
	$\tilde{b_0}$	8.0586	8.1399	8.7348
M	$\tilde{b_1}$	0.0547	0.1140	0.1211
	$\tilde{b_2}$	0.0004	-0.0013	-0.0021
	$\tilde{b_0}$	8.1418	8.1258	8.8870
H	$\tilde{b_1}$	0.0486	0.1383	0.1603
	$\tilde{b_2}$	0.0008	-0.0017	-0.0011

Note: L, M, and H are low, medium, and high-educated men. (Referred in main text section 5.1)

Table A.3: Implied labor elasticity

Wage Change	-20%	-10%	-5%	-1%	+1%	+5%	+10%	+20%	Average
Cohort 50-59									
Own Elasticity	0.57	0.55	0.56	0.52	0.45	0.48	0.45	0.43	0.50
Cross Elasticity	-0.28	-0.29	-0.29	-0.24	-0.25	-0.32	-0.32	-0.33	-0.29
Cohort 60-69									
Own Elasticity	0.33	0.32	0.28	0.31	0.56	0.38	0.37	0.27	0.35
Cross Elasticity	-0.17	-0.18	-0.20	-0.15	-0.13	-0.12	-0.13	-0.15	-0.15
Cohort 70-79									
Own Elasticity	0.20	0.33	0.51	0.70	0.69	1.02	0.89	0.53	0.61
Cross Elasticity	-0.56	-1.02	-1.43	-0.96	-0.97	-0.56	-0.37	-0.22	-0.76

Note: This table shows the implied labor elasticity calculated by changing wage levels. When calculating changes in FLPF, I only include FLFP which is lower than 95% in the baseline model: If FLFP is already very high, it may not be able to respond to changes in wages. For example, if FLFP is 98%, it may increase by the same level given a 1% or 20% increase in women's wages. (Refer to the main text section 5.2)

Table A.4: Robust check of the order of channels (low-educated, cohort 50-70)

Channel	Order	Effect	Order	Effect	Order	Effect
assortative marriage	1	1.14%	1	1.14%	2	1.21%
fertility	2	0.80%	3	1.68%	1	0.73%
childcare cost	3	-4.36%	2	-5.24%	3	-4.36%
Channel	Order	Effect	Order	Effect	Order	Effect
assortative marriage	3	1.14%	2	1.03%	3	1.14%
fertility	1	0.73%	3	1.68%	2	1.58%
childcare cost	2	-4.29%	1	-5.13%	1	-5.13%

Note: This table shows the effect of each family-related channel when they enter the counterfactual study in different orders. (Refer to the main text section 5.3)

B Data Cleaning and Variable Construction

B.1 Family Information Cleaning

I use the following process to detect and correct errors related to family structures in the UHS data:

1. Correct multiple household heads or spouses: I assign correct relation according to the member's age and gender to make sure there is at most one household head and spouse in each family.

2. Correct same-sex marriage: As homosexual marriage is not recognized in China, same-sex marriage in data is mostly due to wrong records. If the household head and the spouse have the same gender, I assign a different sex for the spouse.

3. Impute spouse status: It is necessary to identify if a woman has a spouse. The UHS survey asks for detailed marital status (never married, in marriage, divorced, widowed) after the year 2002 and I can identify if a woman has a spouse by whether she is in a marriage. Before the year 2002, I impute the spouse status based on if a spouse is present in the family. One concern is that the spouse may not be present in the family during the survey. Therefore I compare the imputed status with surveyed data after 2002 to test the accuracy of the imputation method.

Table B.1 shows that, if a spouse is not present in the family during the survey, very likely (87%) the respondent is not in a marriage. If a spouse is present in the family, the respondent is almost certain in a marriage. Therefore we can use the imputed spouse to identify women with spouses. (Referred in main text section 3.1)

Another concern is that we can only impute spouse status if the respondent is the house head or the house head's spouse. I cannot impute spouse status for the house head's parents or children. Table B.2 shows that, if the sample is limited to people has a spouse, 40% of the sample is not used. However, since we are interested in the working-age population and they tend to be the house head. Therefore, 85% of the sample of working age can still be used.

Table B.1: Spouse Imputation Accuracy (based on the year 2002-2009)

	Imputed spouse: no	Imputed spouse: yes
Not in marriage	17,675	299
In marriage	2,535	489,266
Imputation accuracy	87%	99%

Table B.2: Sample Attrition due to Spouse Status

	Working-age		All-age	
	all	with spouse	all	with spouse
Male	315,535	266,039	600,662	367,616
Female	337,643	287,219	611,187	367,913
Total	653,178	553,258	1,211,849	735,529
Remaining Sample	85%		61%	

B.2 Reconcile Two UHS Data

Available UHS data from 2010-2014 only includes 4 provinces: Guangdong, Liaoning, Shanghai, and Sichuan. The first three of them are in the eastern region and are more developed. People in this sample may have higher wages or LFP than the national sample (16 provinces sample). Therefore, I extract a sub-sample of the same 4 provinces from the national sample and compare it with the national sample. I use their difference between

2005-2009 to adjust the sample of 4 provinces.

Figure B.1 is an example to adjust the FLFP (age 25-54). The FLFP of the 4 provinces is 0.8% higher than the national sample between 2005-2009. Therefore, I lower the FLFP of the 4 provinces by 0.8% between 2010-2014 to represent a national sample. Other moments are reconciled in a similar way. This approach is simpler than a complicated re-weight approach and gives a similar result. (Referred in main text section 3.1)

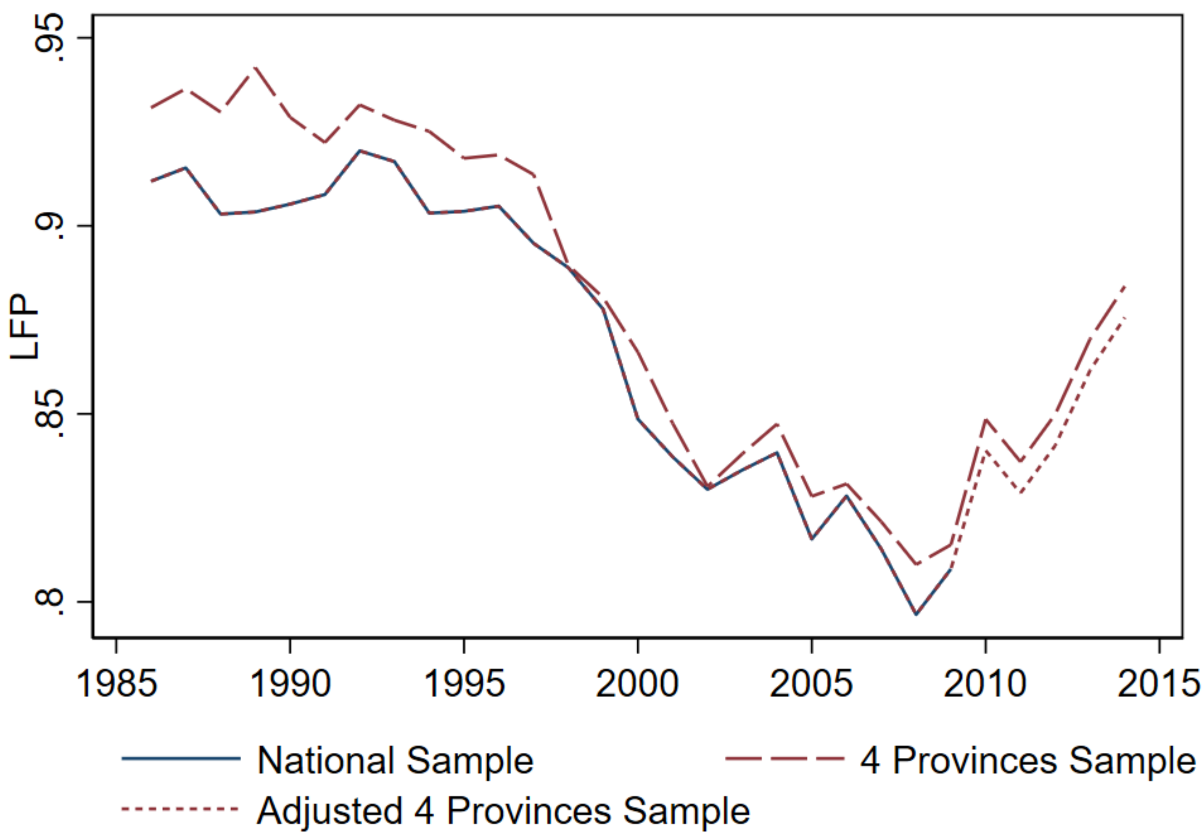


Figure B.1: Reconcile FLFP between Two UHS Data.

Note: the red long dashed line is the 4 provinces sample. The blue solid line is the national sample (16 provinces). The blue short dashed line is the adjusted 4 provinces sample.

B.3 Increased FLFP since 2010

The increased FLFP since 2010 is mostly due to a higher FLFP among women aged 50-54. Figure B.2 compares the FLFP of women aged 45-49 and 50-54 from 1986 to 2014 with the common 4 provinces of the two UHS data. We can see that there is a significant surge in LFP of women aged 50-54 since 2010.

A simple accounting shows the effect of this delayed retirement: The LFP of women aged 25-54 is 81.5% in 2009 and 88.4% in 2014. If we let the LFP of women aged 50-54 in 2014 (74.2%) have the same LFP as in 2010 (35.9%), the counterfactual LFP of women aged 25-54 in 2014 would be 81.0%, keeping the age structural and LFP in other age groups the same, explaining almost all increased FLFP. (Referred in main text section 3.2)

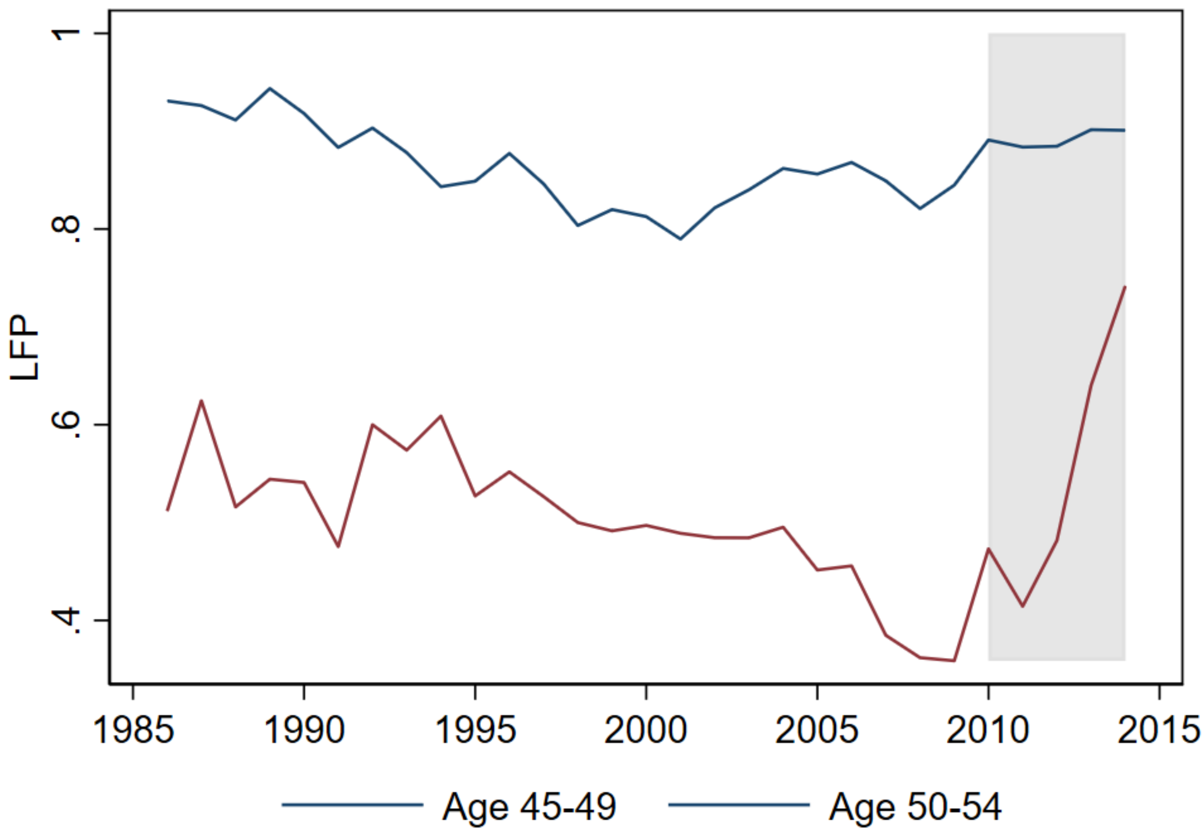


Figure B.2: LFP of Women Age 45-49 and 50-54.

Note: LFP of women aged 50-54 has increased significantly since 2010.

B.4 Adjust for SOE Layoff

Figure B.3 shows the difference in LFP between men and women aged 25-54. There is a kink between 1999 and 2000, which coincides with the SOE layoff period 1998-2004

(Tian, Gong, and Zhai, 2022). I attribute this gap to the SOE layoff and calculate FLFP if there is no such shock (the red dashed line). Theoretically, I can make this adjustment to the FLFP for each age and each cohort. However, the data is not fine enough for accurate adjustment and such adjustment would only increase FLFP by 0.01-0.05% for certain ages. (Referred in main text section 3.2)

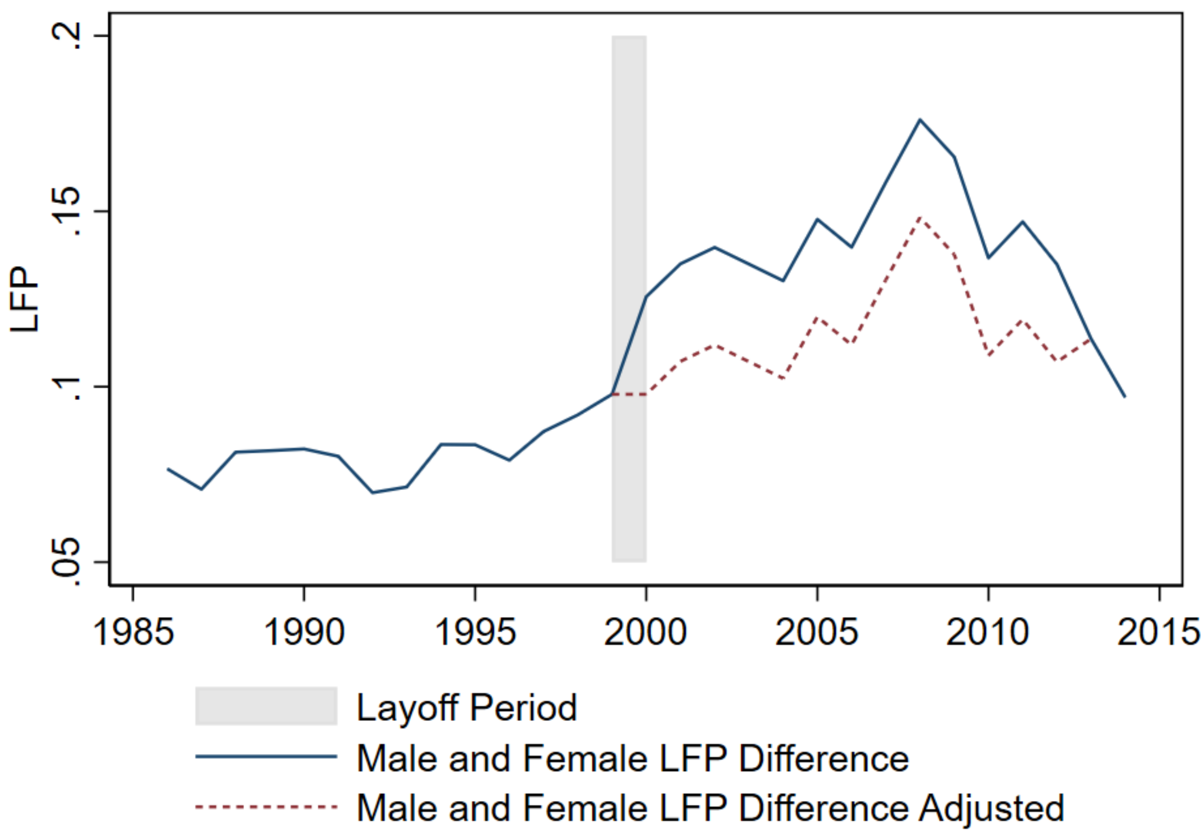


Figure B.3: Adjust FLFP due to the SOE layoff.

B.5 Measure the Degree of Assortative Marriage

The separable Extreme Value (SEV) index is proposed by Chiappori, Costa Dias, and Meghir (2020). To construct the SEV index, consider a simple model:

F\M	E1 (n)	E2 (1-n)
E1 (m)	*r*	*m-r*
E2 (1-m)	*n-r*	*1-n-m+r*

Figure B.4: 2×2 matching table, characterized by (m, n, r).

F and M stand for female and male. E1 and E2 are two levels of education. m, n, r are the portions of each catalog and the total mass is 1. SEV index is:

$$I_{SEV} = \ln\left[\frac{r(1+r-m-n)}{(n-r)(m-r)}\right] \tag{B.1}$$

(Referred in main text section 3.5)