# Explaining the Fall in Female Labor Supply in Urban China

Pengzhan Qian *

Queen Mary University of London

May 2023

**Abstract**: From 1989 to 2009, female labor force participation (FLFP) in urban China decreased by nearly 10 percentage points. This trend is driven by a significant fall in the participation of women without college education. This paper investigates the quantitative importance of multiple channels to explain the fall in FLFP. It features a discrete choice model of labor participation and studies women in cohorts 1950, 1960, and 1970. The counterfactual studies show that a widening gender pay gap can explain 50-80% of the decline in the participation of women without college education. Increased childcare cost accounts for most of the remaining decline, while increased assortative marriage has heterogeneous effects across education levels.

**Keywords:** Female labor supply; gender gap; household; China; Structural Estimation.

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

*Acknowledgement: I thank the guide of my supervisors Giulio Fella and Yu Zheng. I thank the suggestions of Tim Lee, Anna Raute, Vincent Sterk, Wenchao Jin, Elke Jahn, and many others. I also thank participants and comments from QMUL internal seminars, DSE Summer School 2021, DIAL Final Conference, Austrian Labor Market Workshop, EWMES 2022, and RES 2023.

# 1 Introduction

Female labor force participation (henceforth FLFP) has increased in many countries since World War II with positive and substantial impacts on economy and society. For instance, Hsieh et al. (2019) estimate that 20-40% of the economic growth in the U.S. between 1960 and 2010 can be attributed to the increased FLFP. In contrast to this well-known trend, the FLFP of women aged 25-54 in urban China decreased by nearly 10 percentage points from 1989 to 2009 starting from a high level, as illustrated in Figure 1.1.

This pattern is not only significant in scale but also uncommon. First, I do not find similar patterns in typical Western countries, former socialist countries, or other East Asian countries or regions. Second, the FLFP fell when China experienced strong economic growth and the real labor income of women increased significantly, which should have encouraged participation. Third, it cannot be explained by the "U" trend theory of FLFP (Goldin, 1994; Ngai and Olivetti, 2015). This theory predicts that the FLFP will decline when the society moves from agriculture to manufacturing, instead, the FLFP declined in *industrialized* urban China.

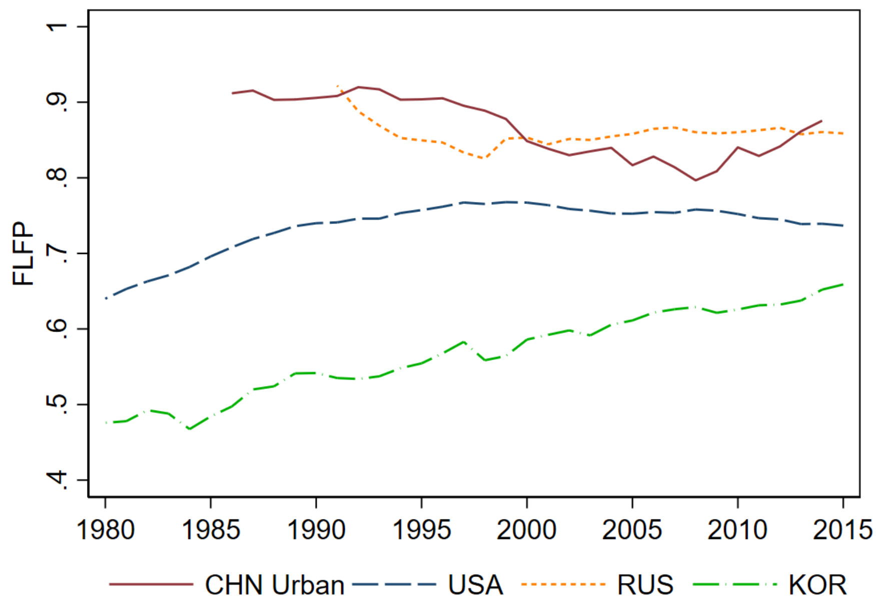



Figure 1.1: Comparison of FLFP (25-54) between Urban China (CHN), the United States (USA), Russia (RUS), and South Korea (KOR).

Note: I use national FLFPs in other countries to cover more years. Urban FLFPs are very close to national FLFPs in these countries. See Figure A.1 in the appendix.

In this paper, I aim to explain the fall of FLFP in urban China. I find that the FLFP declined the most for women without college education, especially for women without senior high school education. To explain this trend, I document several relevant facts. First, despite that women's real labor income increased substantially, their growth rate was lower than that of men's. Second, I provide novel evidence that indicates the cost of child-rearing has nearly doubled over twenty years. These two channels may lower women's participation. Third, the fertility rate decreased only slightly during that period as the "one-child policy" had already been implemented. It may encourage women to work more. Fourth, using the method proposed by Chiappori et al. (2018), I confirm that marriage became more assortative (people with similar education levels tend to marry more) while its effect on FLFP is subjected to empirical analysis.

Despite these potential channels having been examined separately, the complicated interactions among them are not considered in previous studies. Therefore, I build a dynamic life-cycle model of married women to quantify the contribution of each channel.

In the model, a married woman starts life at age 25, with her and her husband's education levels exogenously given. Each year, the woman chooses whether to work after observing her individual productivity shock until retires, while her husband always works. Each year, she may have one child exogenously if she does not have a child before. If her child is younger than age 6, she can either work but pay a childcare fee, or stay at home to take care of her child. The woman has the incentive to work to build up her human capital for a higher future income. However, if her husband's income is much higher than hers, or if the childcare fee is prohibitively high, she would choose not to work.

I use this model to study women's labor participation in cohorts 1950, 1960, and 1970. This model is estimated for each cohort by the simulated method of moments (SMM). The average FLFP and income at each age and each cohort are targeted moments. The model has a good fit and can reproduce these patterns well.

I use counterfactual studies to analyze the importance of each channel. I first study the effect of income by simulating a counterfactual FLFP if men and women of the 1950 cohort can earn as much as the young cohorts. The counterfactual study shows that the faster growth of husbands' incomes has a strong income effect. It pulls down the FLFP of women without senior high school education in the 1960 and 1970 cohorts by 2.6 percentage points and 7.3 percentage points respectively, explaining 50-80% their difference in FLFP. For women with senior high school education, the widening gender pay gap is also the main driving force.

I then analyze the effects of changes in childcare cost, marriage, and fertility rate jointly with changes in income. I find that a higher childcare cost reduces the FLFP by 2-4 percentage points per year for all women, explaining a significant share of the decline in FLFP. An increased assortative marriage makes women without senior high school education work more but makes women with senior high school education work less. A lower fertility rate increases the FLFP slightly. Finally, I connect the model findings to the decline in FLFP over the years. The disparity between the 1950 and 1970 cohorts accounts for 36-48% of the reduction in FLFP from 1989 to 2009.

These exercises illustrate that a widening earning gap within a family can result in a significant decline in the FLFP. Family related-channels can affect the FLFP by influencing the earning gap within a household. Addressing income inequality should be the primary focus of policy interventions that aim to promote the FLFP.

**Related Literature** This paper contributes to the literature on female labor supply in three ways. First, the increased FLFP in the U.S. and U.K. has attracted considerable research interests as early as Heckman and Macurdy (1980). The monograph of Goldin (2021) documents the changes in FLFP across cohorts in detail. From a macro perspective, Acemoglu et al. (2004) link the increased FLFP with production demand during World War II. Ngai and Olivetti (2015) demonstrate the role of structural change in boosting FLFP. From a household perspective, Attanasio et al. (2008) analyze the effects of education, the gender pay gap, and childcare costs on the increased FLFP. Chiappori et al. (2018) extend the analysis of FLFP by endogenizing education and family choices. These papers focus mostly on the female labor supply in the U.S. and U.K. while not all countries, even not all industrialized counties experience increased FLFP (Olivetti and Petrongolo, 2016). This paper fills this gap by studying the substantial fall in FLFP in China, one of the most populated countries.

Second, the paper contributes to the literature on gender inequality by demonstrating the effect of a widening gender pay gap empirically. The literature establishes that a narrowing gender pay gap is a dominant factor to explain the increased FLFP in the U.S. (Eckstein and Lifshitz, 2011) and the U.K. (Blundell et al., 2016). I employ China's unique gender pay gap trend to study the classic topic of female labor supply from a novel direction, showing a widening gender pay gap could depress FLFP. This finding echoes recent studies that focus on the slowdown of participation rate convergence or declined male participation rate in the U.S. Albanesi and Prados (2022) argue that the slowing convergence of FLFP among college-educated women is due to faster growth of husbands' wages. Coglianese (2018) shows that the growing wages of wives partially explain the declined participation of men.

Third, this paper is among the few to use a structural model to study the female labor supply in China. Keane and Wolpin (2009) and Keane (2011) highlight the value of using structural models in studying female labor supply but this approach is less adopted in studies on China. Jin (2016) and Gao (2020) use structural models to study the labor supply of near-retirement women and pension policies. Román (2022) study how the unbalanced sex ratio can explain the decline in females' working hours. To the best of my knowledge, this paper is the first to study the declining FLFP among *working-age* women using a structural model. The decline of FLFP in China attracts considerable academic attention (Feng et al., 2017): Hare (2016) studies the effect of women's own wages; Song and Dong (2018) studies the effect of increased childcare cost; Feng and Tang (2019) study the effect of assortative marriage. This paper bridges multiple channels in a structural model to determine the importance of each channel quantitatively.

The rest of the paper is organized as follows. Section 2 describes several important facts and data sources. Section 3 presents the model in detail and Section 4 shows the estimation results, model validity, and uses counterfactual studies to determine the quantitative importance of each channel. Section 5 concludes the paper with further discussion.

# 2 Data and Facts

## 2.1 Data and related background

The main microdata used for this study is the Urban Household Survey (UHS). UHS is a repeated cross-section data collected by the National Bureau of Statistics of China. The data covers the period from 1986 to 2014. Between 1986 and 2009, it covers 16 provinces, 390,000 households, and 1.2 million individuals. Only data from four provinces are available between 2010 and 2014, covering 21,000 households and 0.16 million individuals. I detail the process of data cleaning in Appendix B.

In this study, the labor supply of married women aged 25-54 is the main focus of this study. By the age of 25, most people have finished their education. Therefore, the aggregate FLFP is less likely to be influenced by a longer education. As the official retirement age is 50-55 for women (and 60 for men) depending on occupations, by the age of 55, most women have retired.

Marital status was not directly surveyed until 2002. Instead, the survey asked about the relationship between one family member and the household head. Therefore, it is possible to infer that a woman is married if she is the spouse of the household head or if she is the household head and has a spouse. This definition excludes women whose husbands were not present during the survey and married women who cohabit with their parents, as they are recorded as ”child” or ”child-in-law”. However, according to this definition, more than 84% of all women aged 25-54 could be identified as married. I do not focus on single women for two reasons. First, it is not easy to identify single women before 2002. Second, after 2002, single women can be identified while they have higher FLFP than married women (Figure A.2 in the appendix ). As more women are remains single, it should *increase* the aggregate FLFP.

To investigate FLFP by education groups, I define low-educated as people with junior-high-school degrees or below (less than 9 years of education). Medium-educated are people with senior-high-school degrees or equivalent (10-12 years of education). High-educated are people with some college education or more (more than 13 years of education).

Although women's education has improved significantly, low and medium-educated women still account for a major share in urban China. To capture the change in educational levels, I focus on young women aged 25-29. In 1987, only 2% of young women aged 25-29 have some college education or more. In 2010, more than 30% of women aged 25-29 are high-educated but two third of women aged 25-29 are still low and medium-educated (Table 2.1). Therefore, the labor market outcome of non-college educated women is still worth studying.

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

Table 2.1: Share of women in each education group in urban China (age 25-29).

| Education | Definition | 1987 | 1995 | 2000 | 2005 | 2010 |
|---|---|---|---|---|---|---|
| Low | $\leq$ junior high school ($\leq$ 9 years of education) | 65% | 67% | 62% | 55% | 45% |
| Medium | senior high school (10-12 years of education) | 33% | 24% | 25% | 24% | 21% |
| High | $\geq$ college ($\geq$ 13 years of education) | 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.

## 2.2 Fact 1: Changes in FLFP

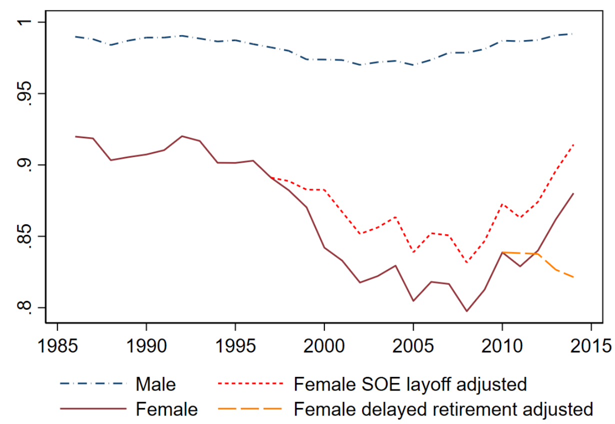
Figure 2.1a shows trends of LFP by gender across years for people aged between 25 and 54. From 1989 to 2009, men's LFP falls by just less than 1 percentage point while women's LFP dropped by nearly 10 percentage points. Since 2010, the FLFP has increased but it is due to delayed retirement. The red dashed line adjusts the FLFP by assuming the FLFP of women aged 50-54 is kept at the 2010 level.

Like many former socialist countries with centrally planned economies, China has experienced a transition period, which is often associated with a decline in FLFP. Around the year 2000, 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 dashed line in Figure 2.1a adjusts this shock by assuming the excess decrease in FLFP between 1998 and 2000 was due to this layoff. Still, the FLFP has dropped by nearly 6 percentage points under this strong adjustment. See Appendix B.4 for details of these two adjustments.

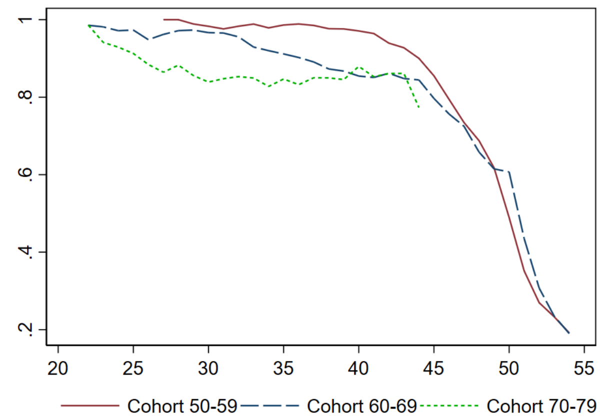
Figure 2.1b, 2.1c, and 2.1d break the FLFP by cohorts (cohort 1950-1969, cohort 1960-1969, cohort 1970-1979) and education groups. It is clear that the decline of FLFP is driven by low and medium-educated women while the LFP of high-educated women remains high. For the low-educated group, their LFP dropped 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 FLFP after the year 2009 in Figure 2.1a.

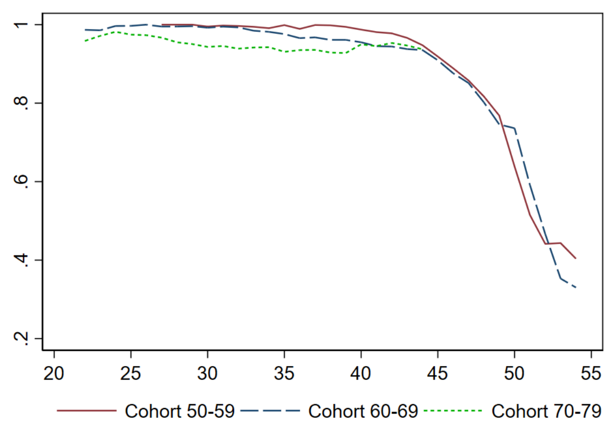
We can also notice year effects in the cohort view. The decline of FLFP for cohort 1960 is from age 30-40, roughly corresponding to the year 1995-2005. The decline of FLFP for cohort 1970 is from age 20 to 30, also fell between the years 1995 to 2005.



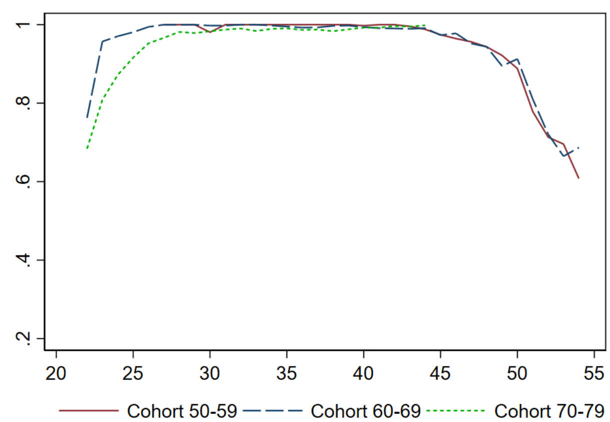

(a) LFP by gender across years




(b) FLFP of low-educated




(c) FLFP of medium-educated




(d) FLFP of high-educated

Figure 2.1: FLFP by years and by cohorts.

Note: FLFP is driven down by low and medium-educated women. The orange dotted line in Figure 2.1a adjusts the FLFP for the SOE layoff. See Appendix B.4 for details. FLFP is not further adjusted in the cohort views as such adjustment would just increase FLFP by 0.01-0.05 percentage points to certain ages.

### 2.2.1 Decomposition of the FLFP

Graphic evidence suggests that the decline of the FLFP of low and medium-educated women plays an important role in driving down the aggregate FLFP. However, the decline in the aggregate 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 aggregate FLFP explained by changes in the FLFP of certain groups, keeping demographic structures the same.

The first column of Table 2.2 shows that, between 1989 and 2009, the aggregate FLFP decreased by 9.5 percentage points. The second column shows that, if the age structure holds the same, the change in FLFP of each age group would have lowered the aggregate FLFP by 6.4 percentage points, explaining 67% of the decline. The fourth column shows that, if the age and education structure holds the same, the change in FLFP of each age and education group would have lowered the aggregate FLFP by 11.3 percentage points, meaning that the composition effect of education has *increased* the aggregate FLFP. Similar patterns could be found if the whole period is separated into two decades.

Combined with graphic evidence, we can confirm that the declining aggregate FLFP in urban China is mostly due to the decline of FLFP within low and medium-educated women, rather than composition effects.

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

Table 2.2: 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 FLFP or demographic structure is in the parenthesis.

## 2.3 Fact 2: Gender Pay Gap

Labor income has grown substantially in China, especially after the economic transition around the year 2000. However, the growth of men's income is higher than that of women's (Gustafsson and Li, 2000; Chi and Li, 2008; Ge and Yang, 2014). Figure 2.2 illustrates the average annual labor income of low and medium-educated men and women. The annual labor 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).

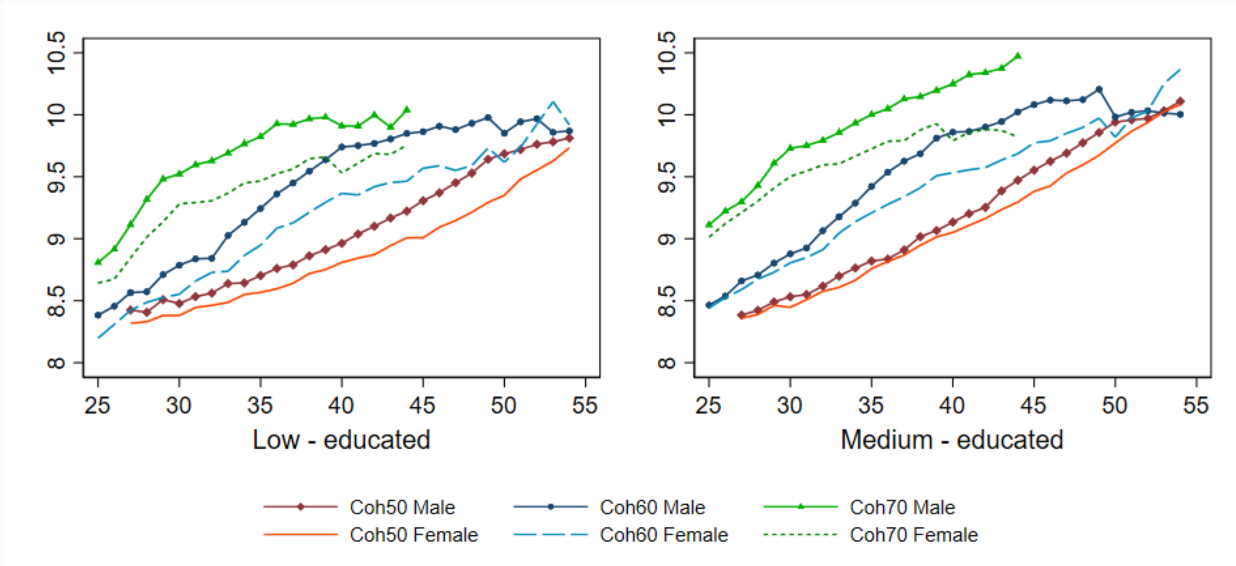



Figure 2.2: Logarithm of real income of low and medium-educated people (age 25-54, in 2009 price).

The left figure in Figure 2.3 shows the ratio between men's and women's annual income after controlling for their experience. We can see that the pay gap is widening for all education groups since 1990, especially between 1995 and 2005. These trends are robust when industry and province fixed effects are further controlled or measured in an alternative approach (Figure A.3 in the appendix). Admittedly, the annual income is not the perfect measure for the gender pay gap as the working hours of men and women may be different. Hours are only surveyed between 2002-2006 in the UHS survey and for these years, a similar trend of the widening gender pay gap could also be found.

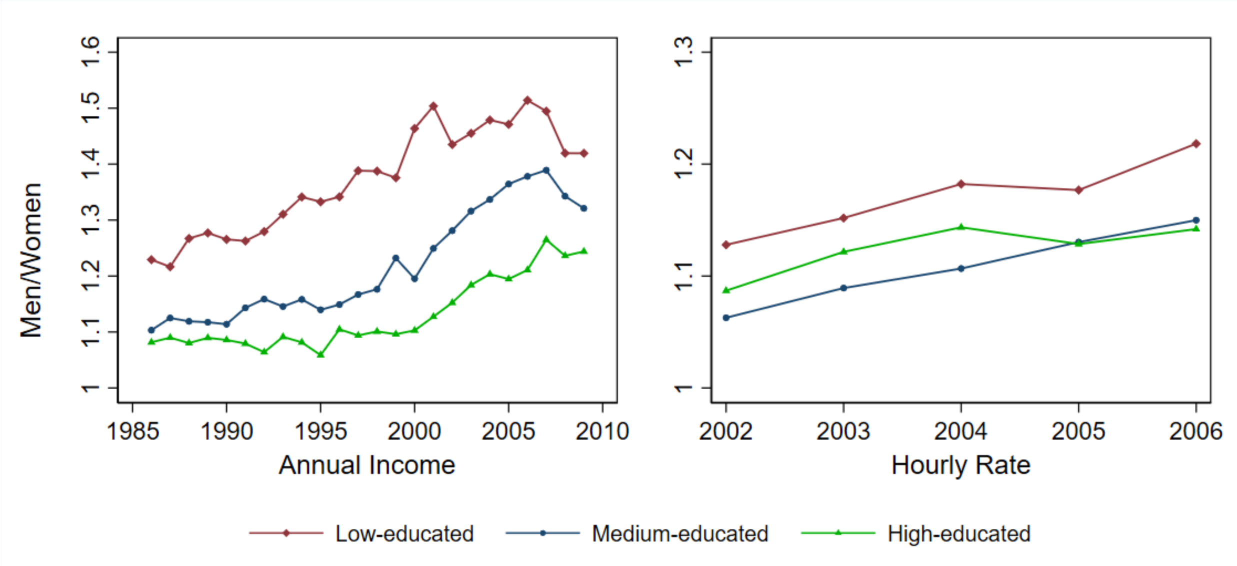



Figure 2.3: Gender pay gap (men/women) across years by education groups (age 25-54).

Note: The hourly rate is only available between 2002 and 2006.

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) (Stafford, 2015). $p_{it}$ is the participation of household $i$ in year $t$. $\widehat{w_{it}}$ is the woman's potential earning and $\widetilde{w_{it}}$ is the man's earning. $Z_{it}$ is other control variables including age, the square of age, and the presence of children younger than age 6. $FE$ are year, city, cohort, and education fixed effects. As I do not have panel data therefore I cannot control for household fixed effect but I can year, age, and cohort together. Here I do not use logarithm transform as it would be easier to interpret elasticity.

I use a Heckman regression to predict the potential wages for women as Equation 3. $Z_{it}$ and $FE$ are the same as in Equation (2). $\lambda_P()$ is the control function to approximate the inverse Mills ratio. I use men's earnings as an instrument to predict the participation of women. The results for the selection and wage regression are shown in columns (1) and (2) in Table 2.3 respectively. The fitted value of Equation 3 is the women's potential wage, which is used to estimate Equation (2). The results are shown in column (3) and 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. That is, an increase in women's *own* earnings would encourage them to work more. However, an increase in their husbands' earnings would discourage them to work.

These elasticities are estimated from the data directly. To the best of my knowledge, there is no relevant policy shock available in China's setting to achieve better 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 an aggregate level, regressions also show that the city-level gender wage gap is negatively related to local FLFP (Tabel A.1 in the appendix).

$$\text{Probit}(p_{it}{=}1) = \beta_1 \widehat{w_{it}} + \beta_2 \widetilde{w_{it}} + \gamma Z_{it} + FE \tag{2}$$

$$w_{it} = \gamma_1 Z_{it} + \lambda_P(\beta_1 \widetilde{w_{it}} + \gamma_2 Z_{it} + FE) + FE \tag{3}$$

Table 2.3: Household-level Wage and FLFP

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Selection | Wage | Probit(p=1) | Elasticity in (3) |
| $\widehat{w}$ | | | 0.000*** | 0.263*** |
| | | | (0.000) | (0.003) |
| $\widetilde{w}$ | -0.000*** | | -0.000*** | -0.032*** |
| | (0.000) | | (0.000) | (0.002) |
| age | 0.316*** | 892*** | 0.286*** | |
| | (0.006) | (54.3) | (0.008) | |
| age$^2$ | -0.004*** | -11.0*** | -0.004*** | |
| | (0.000) | (0.693) | 0.000 | |
| child | -0.203*** | -359*** | -0.137*** | |
| | (0.013) | (78.3) | (0.016) | |
| inverse mills ratio | | 6972 | | |
| | | (252) | | |
| Observation | 258,579 | 258,579 | 257,681 | |
| Year FE | Yes | Yes | Yes | |
| City FE | Yes | Yes | Yes | |
| Cohort 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.

## 2.4 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. After age 6, children can go to the primary school. I find the childcare cost as a ratio to household expenditure has increased from the early 1990s to the early 2000s in Figure 2.4a. 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 2.4b), 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 required most families to have at most one child. This policy was implemented gradually in the early 1980s and lifted in 2015. Therefore, some women in the cohort 1950 were not affected and may have more than one child.


(a) Childcare cost/Household expenditure.

(b) Probability of giving birth.

Figure 2.4: Children-related trends.

Note: The childcare care trend is only available until 2012 as the household consumption data in 2013 and 2014 are not comparable with other years.

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.

## 2.5 Fact 4: Increased Assortative Marriage

The fourth fact I looked into is increased assortative marriage. Assortative marriage means people marry others with similar characteristics. This study focuses on sorting by 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, it is necessary to take the change in the distribution of education into account. For example, the probability of having a low-educated husband given the woman is also low-educated, only increases from 86% to 90% from cohort 1950 to cohort 1970. From this perspective, the marriage market faced by low-educated women does not change significantly. However, the proportion of low-educated men decreased from 45% to 24% across cohorts. In other words, low-educated women are marrying slightly more low-educated men, given that the supply of low-educated men has halved, indicating a much more assortative marriage market.

Chiappori, Costa Dias, and Meghir (2020) propose using the Separable Extreme Value index (SEV) to measure the degree of assortative marriage. The SEV is calculated as the logarithm of the ratio between the number of assortative marriages and non-assortative marriages. It is based on a transferable utility model and can account for changes in the population's education, making it a more robust indicator compared to others. With this measurement, I find that marriage is becoming more assortative for all education groups, especially for low-educated women. A higher value of the SEV index indicates a higher degree of assortative marriage in Table 2.4). I restrict the sample to women aged 30-50 to mitigate the difference in marriage age and death risks across education groups. However, the degree of assortative marriage remains similar when no age restriction is imposed (Table in the appendix).

Table 2.4: Degree of assortative marriage (SEV index).

| Cohort | Low-educated | Medium-educated | High-educated |
|---|---|---|---|
| 1950-1959 | 3.73 | 3.06 | 3.94 |
| 1960-1969 | 4.22 | 3.08 | 4.09 |
| 1970-1979 | 5.03 | 3.59 | 4.49 |
| $\Delta_{70-50}$ | 35% | 18% | 14% |

Note: SEV index is calculated based on women aged 30-50. Higher SEV values mean higher degrees of assortative marriage. See Appendix B.5 for details.

# 3 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 FLFP 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 2.1, the LFP of low and medium-educated women are almost equally high between 22 and 25.

In the model, the gender pay of men and women is regarded as exogenous. I discuss potential mechanisms and puzzles in Section 5.

## 3.1 Outline of the Model

A woman of cohort $\theta$ enters the economy at age 22 with a given education level $e \in \{L, M\}$. She lives in a family with a man of the same age [1]. Her husband's education

[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.2).

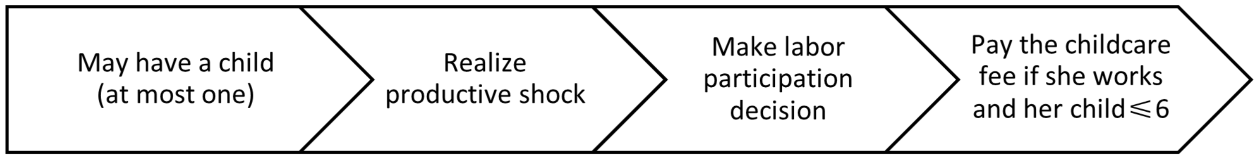



Figure 3.1: Model timeline of a year.

level $\tilde{e} \in \{L, M, H\}$ (all variables related to the husband are distinguished by "$\sim$") is determined by the cohort-specific matching matrix (which embodies the degree of assortative marriage). The family does not dissolve.

The woman's age is denoted by $a$ and she maximizes her own lifetime utility by making labor decisions $p$ in every period after observing her personal productivity shock. She retires at 55 and left the model.

In each period, a childless woman has a cohort-specific probability 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 $\chi$. The family could pay childcare cost $\kappa_{\theta+a}$ ($\kappa$ 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 3.1.

## 3.2 Preference and Budget Constraint

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

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

$C_a$ 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. $\rho$ is the risk aversion coefficient. $\gamma_1$ and $\gamma_2$ 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:

$$V_a^\theta(Z_a) = \max_{\{p_\tau\}_{\tau=a,\ldots,54}} E\left\{ \sum_{\tau=a}^{54} \beta^{\tau-a} u(C_a, p_a; n_a, e) | Z_a \right\} \tag{5}$$

under the budget constraint:

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

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

### 3.2.1 Income Process

The annual labor 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 $a$. $\epsilon_a$ is the stochastic shock.

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 year of experience. Otherwise, her experience depreciates by $\delta$.

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

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 the i.i.d. error term or the variance of the ability to draw $\epsilon_a \sim N(0, \sigma)$.

The husband's labor income is also determined by a Mincer-type equation. I assume the husband always works (therefore his experience is equal to his age) and there is no variance in income to simplify the problem:

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

# 4 Estimation, Results, and Counterfactual Studies

## 4.1 Estimation and Results

$\rho$ and $\beta$ are set as 1.5 and 0.98 as in the literature (Attanasio et al., 2008). Parameters in husbands' income function ($\tilde{b}$) are estimated from the data by Equation (9) and the coefficients are listed in Table A.2 in the appendix. The degrees of assortative marriage, fertility rates, and childcare costs are calculated from the UHS data or the census data.

The remaining parameters are estimated by 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 ($\gamma_1, \gamma_2$) 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 average FLFP and annual labor 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 $k$th moment for simulation. $M^d$ is the moment in data and $M^s(\Theta)$ is the simulated moment with parameter $\Theta$. The weighting matrix is the inverse matrix of the variance of data moments following the common approach of SMM (Eckstein and Lifshitz, 2011; Blundell et al., 2016). The asymptotically optimal weighting matrix is not used as Altonji and Segal (1996) criticize its potentially poor small-sample properties.

Table 4.1 summarizes the parameters and their estimations in the model.

Table 4.1: Summary of Parameters

| External Estimated | Values or Sources |
|---|---|
| Risk Aversion Parameter ($\rho$) | 1.5 (Attanasio et al., 2008) |
| Discount Factor ($\beta$) | 0.98 (Attanasio et al., 2008) |
| Men's income parameters ($\tilde{b_0}$, $\tilde{b_1}$, $\tilde{b_2}$) | Data |
| Assortative marriage, Fertility rate, Childcare cost | Data |
| **Internal Estimated** | **Target Moment** |
| Women's income parameters ($b_0$, $b_1$, $b_2$) | |
| Depreciation rate ($\delta$) | FLFP and Wage by cohort and age |
| Variance in potential wages ($\sigma$) | |
| Preference parameters ($\gamma_1$, $\gamma_2$) | Jointly estimated across cohorts |

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). The estimation results for low and medium-educated are listed in Table 4.2 and 4.3 respectively. $\chi^2$ is the Pearson cumulative test statistic and it is smaller than the corresponding critical value in all cases. Therefore models are not rejected in over-identification 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 4.1. 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.

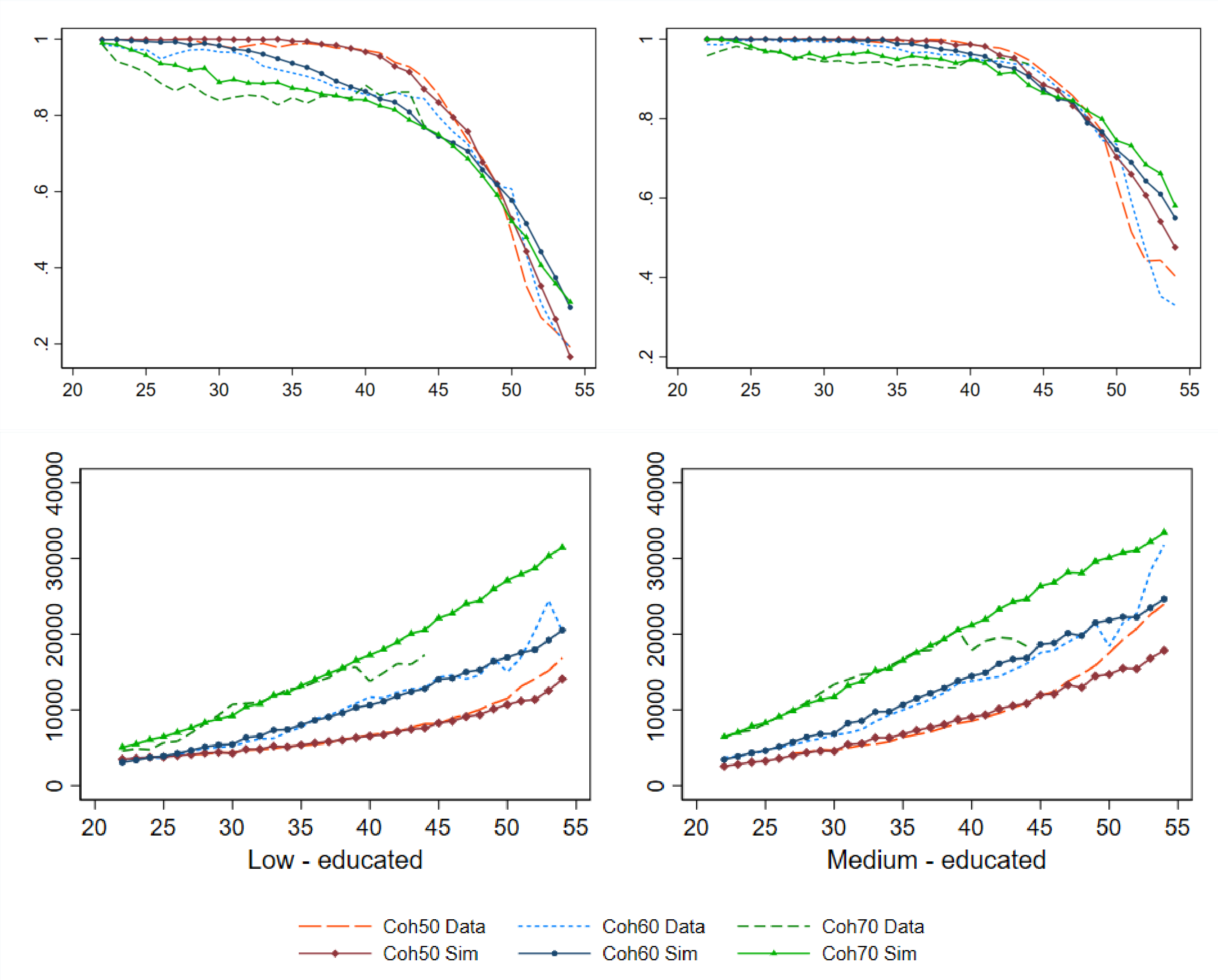



Figure 4.1: Fit of low and medium-educated women's LFP and income (in 2009 price).

## 4.2 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 4.2: 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) |
| $\delta$ | 0.167 | 0.150 | 0.019 |
| | (0.025) | (0.022) | (0.033) |
| $\sigma$ | 0.471 | 0.492 | 0.317 |
| | (0.007) | (0.009) | (0.016) |
| $\gamma_1$ | | -0.186 | |
| | | (0.002) | |
| $\gamma_2$ | | 0.000 | |
| | | (0.000) | |
| $\chi^2$ | 17.4 | 30.9 | 20.6 |
| $\chi^2_{critical}$ | 66.3 | 77.9 | 54.6 |

Table 4.3: 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) |
| $\delta$ | 0.040 | 0.077 | 0.000 |
| | (0.006) | (0.012) | (0.012) |
| $\sigma$ | 0.639 | 0.608 | 0.387 |
| | (0.005) | (0.007) | (0.012) |
| $\gamma_1$ | | -0.178 | |
| | | (0.003) | |
| $\gamma_2$ | | -0.000 | |
| | | (0.000) | |
| $\chi^2$ | 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. $\delta$: experience depreciation rate. $\sigma$: variance of ability draw. $\gamma_1$, $\gamma_2$: preference parameters (estimated jointly across three cohorts). $\chi^2$ is the Pearson cumulative test statistic and $\chi^2_{critic}$ 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 4.4 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 4.4: 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 (**??**) 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 FLFP for high-educated groups before age 50 (Figure A.4 in the appendix).

## 4.3 Counterfactual Studies

With the model, I could now study the contribution of each channel. I denote the FLFP of cohort 50-59 as $L_{50}$, FLFP of cohort 60-69 as $L_{60}$, and FLFP 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 FLFP $L'_{50}$ (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 FLFP $L''_{50}$.

As there are multiple channels, there are two approaches to study 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 (See Table A.4 in the appendix).

### Counterfactual study for assortative marriage

The counterfactual study for assortative marriage is less straightforward. One simple approach is to substitute cohort 50 with the marriage matrix of younger cohorts. In this way, however, women in cohort 50 just have the same conditional probability to find a husband of certain education level $Prob_{60}(E_{men}|E_{women})$, $Prob_{70}(E_{men}|E_{women})$, rather than facing the same degree of assortative marriage market as younger cohorts.

Another approach is to find the counterfactual conditional probability given the education distribution of cohort 50 and the degree of assortative marriage of younger cohorts $Prob'(E_{men}|E_{women}, \text{Edu}_{50}, \text{SEV}_{60})$, $Prob''(E_{men}|E_{women}, \text{Edu}_{50}, \text{SEV}_{70})$.

### Income Channel

Figure 4.2 shows FLFP 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 FLFP 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, the FLFP 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 FLFP (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 FLFP [3].

The second row of graphs compares the income channel and the family channel. The

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

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 FLFP to be explained by the family channel.

**Familly 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.

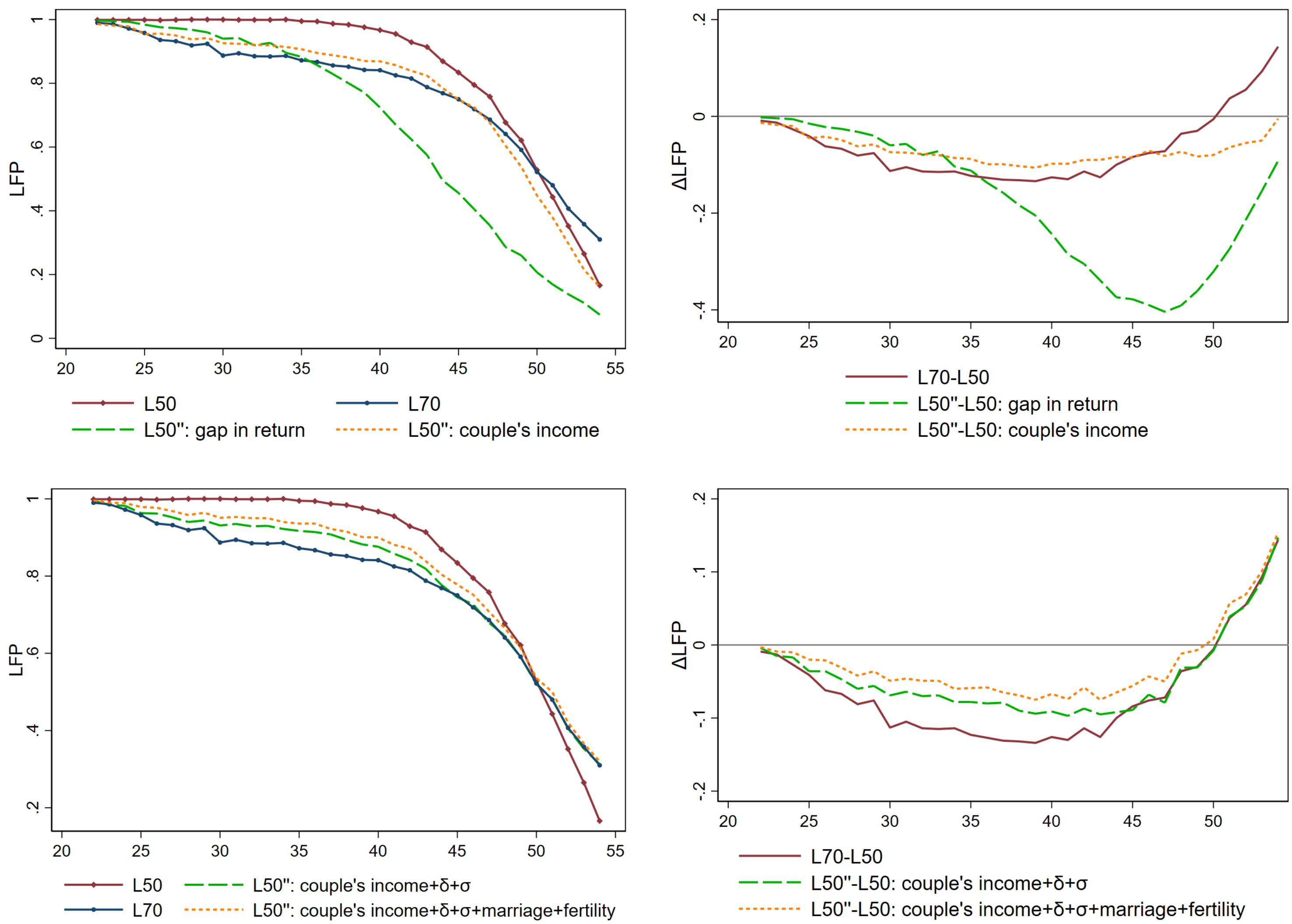



Figure 4.2: Counterfactual FLFP for low-educated women between cohort 1950 and 1970.

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

Figure 4.3 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 FLFP.

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 FLFP). Increased childcare cost also matters for them, which explain almost all of the declined

FLFP on top of the income channel.

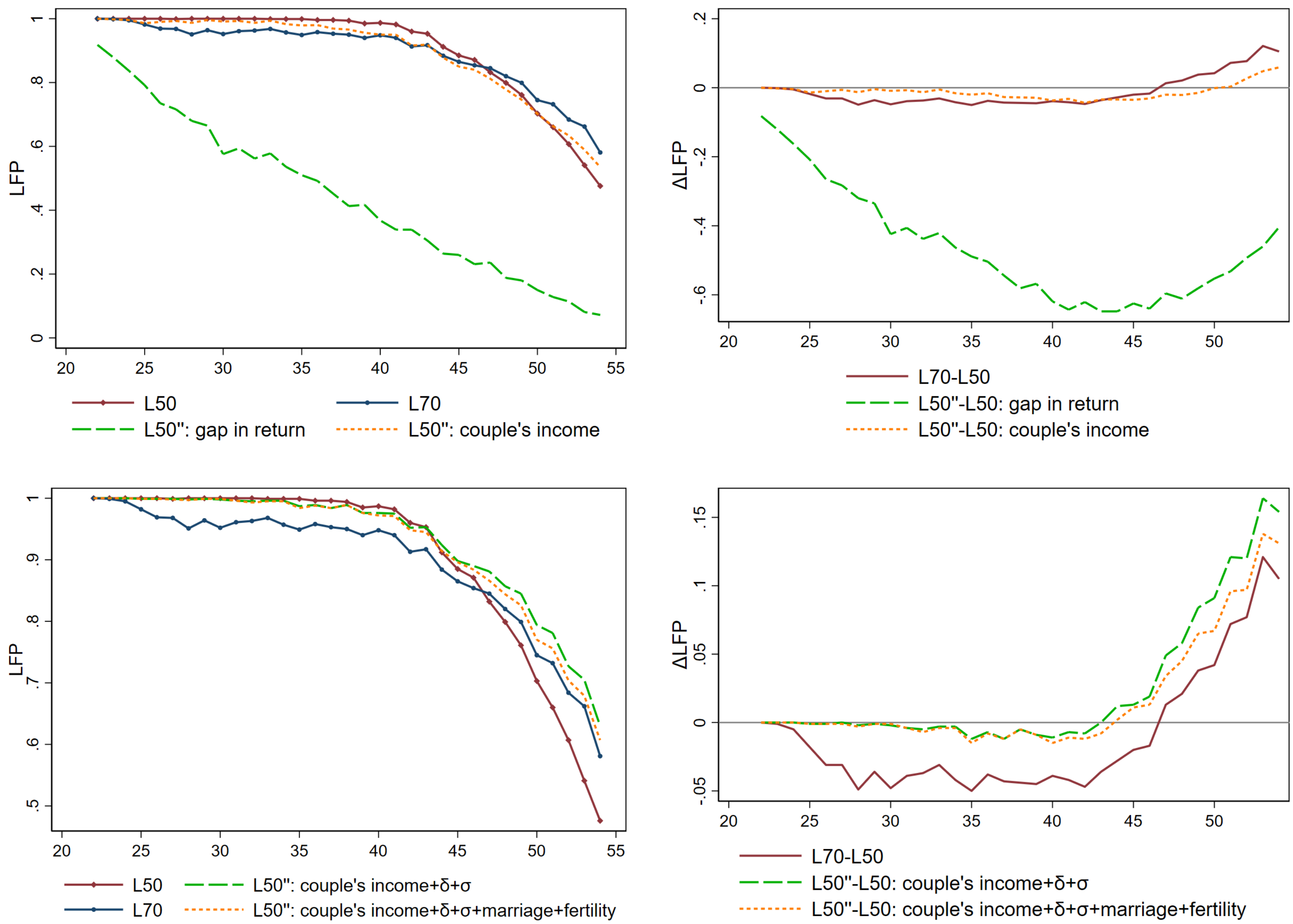



Figure 4.3: Counterfactual FLFP for medium-educated women between cohort 1950 and 1970.

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

Table 4.5 summarizes quantitative results. The first row of the table shows average changes of FLFP between 22-49 in data. I focus on this period because it is when most FLFP 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 FLFP increases by 1.9% on average. I then let them have the same gap in wage *upon* the same gap in return and their FLFP 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 cohorts 50 and 60, changes in couples' income can explain 50%-80% the declined FLFP. Except for low-educated women between cohorts 50 and 60, changes in the return gap have strong 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 FLFP for all women. Increased childcare costs reduce FLFP for all women except medium-educated women between cohort 50 and 60 and they are the main driving forces in family-related channels.

Table 4.5: Effect of each channel on FLFP 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 $\tilde{b_0}$). $\delta$: depreciation rate. $\sigma$: variance in ability. Family structures combine assortative marriage, fertility rate, and childcare cost.

## 5 Conclusion and Discussion

The declining FLFP in China is a strong and uncommon 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 percentage points each year at a young age for low-educated women in the 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 the aggregate FLFP in 1989 by 5 percentage points, explaining 48% of the decline in the aggregate FLFP between 1989 and 2009. From Table 4.5, 75% of these cohort-level changes are explained by the income channel, which corresponds to 36% of the aggregate decline in FLFP between 1989 and 2009.

The next question would be: why the gender wage gap has been widening? It is a complicated question that goes beyond the scope of this paper. Instead of offering an answer, I discuss some potential explanations. First, discrimination and social norm could widen the pay gap. Although the female participation rate in China is relatively high, it does not mean the welfare system and social support for women are proportionally adequate. For example, gender-specific job advertisements were not banned until recently (Kuhn and Shen, 2012, 2023). Chen and Ge (2018) find that the attitude to work can be inherited within a family. Second, firms' expectations may also play a role. Xiao (2021) shows that if firms expect women to leave the labor market more because of family reasons, they pay women lower wages in Finland. This mechanism could be complementary to the findings in this paper: lower participation leads to a larger pay gap, which results in an even lower participation rate.

However, a strong first-order explanation is still to be found. I find two puzzling trends: First, gender pay gaps widen in both manufacturing and service sectors. This contradicts the structural change theory which predicts women benefit from the structural change as they have comparative advantages in the service sector (Ngai and Petrongolo, 2017). Further analysis shows that pay gaps widen more in capital-intensive industries. One potential explanation is that male workers are more compatible with capital or capital-embodied technology. Second, pay gaps widen in most occupations. The pay gap even widens for government clerks after controlling for experience, while their wages only depend on tenures and ranks. It may suggest that men are promoted faster than women as evidenced by that the majority of high-ranking officers in China are men. These potential channels are worth exploring more to explain the widening pay gap.

# 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.

ALTONJI, J. G. AND L. M. SEGAL (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business Economic Statistics*, 14, 353–366.

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.

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

CHEN, X. AND S. GE (2018): "Social norms and female labor force participation in urban China," *Journal of Comparative Economics*, 46, 966–987.

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.

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*.

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.

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," .

GE, S. AND D. T. YANG (2014): "Changes in China's wage structure," *Journal of the European Economic Association*, 12, 300–336.

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

——— (2021): *Career and family*, Princeton University Press.

GUSTAFSSON, B. AND S. LI (2000): "Economic transformation and the gender earnings gap in urban China," *Journal of Population Economics*, 13, 305–329.

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.

HSIEH, C.-T., E. HURST, C. I. JONES, AND P. J. KLENOW (2019): "The Allocation of Talent and U.S. Economic Growth," *Econometrica*, 87, 1439–1474.

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

KEANE, M. P. (2011): "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 49, 961–1075.

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.

KUHN, P. AND K. SHEN (2012): " Gender Discrimination in Job Ads: Evidence from China *," *The Quarterly Journal of Economics*, 128, 287–336.

——— (2023): "What Happens When Employers Can No Longer Discriminate in Job Ads?" *Amer-*

*ican Economic Review*, 113, 1013–48.

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.

ROMÁN, F. J. R. (2022): "The Sex Ratio, Marriage and Bargaining: A Look at China," .

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

STAFFORD, T. M. (2015): "What Do Fishermen Tell Us That Taxi Drivers Do Not? An Empirical Investigation of Labor Supply," *Journal of Labor Economics*, 33, 683–710.

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.

XIAO, P. (2021): "Wage and Employment Discrimination by Gender in Labor Market Equilibrium," Working Papers 144, VATT Institute for Economic Research.

# A Supplementary Figures and Tables

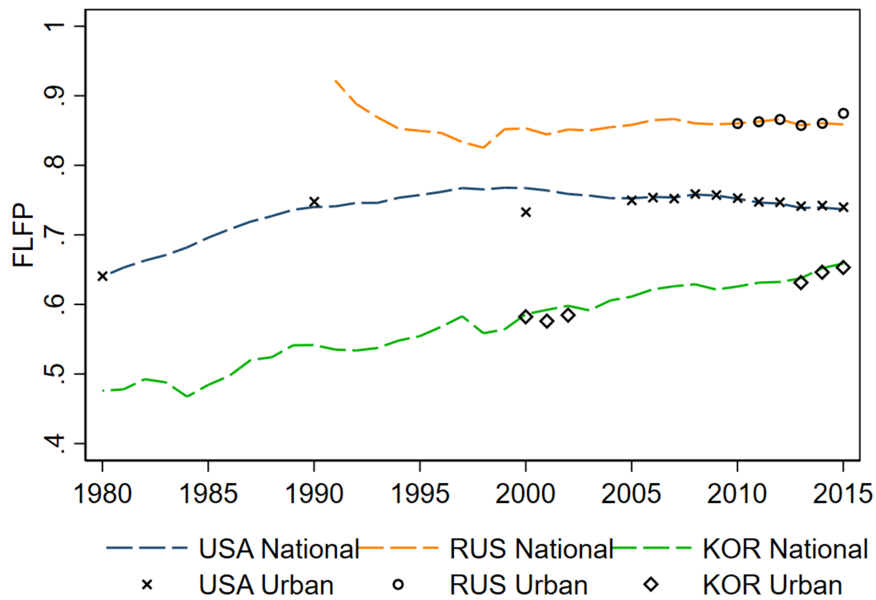
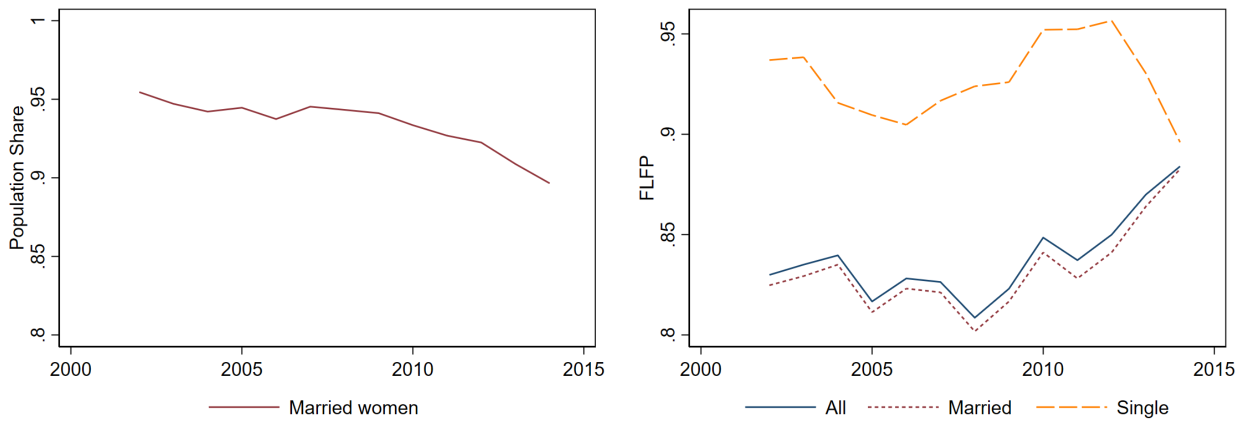



Figure A.1: Urban and national FLFPs in the United States(USA), Russia (RUS), and South Korea (KOR).

Note: Urban FLFPs in these countries are very close to their national FLFPs due to high urbanization rates. Source: International Labour Organization (ILO) and IPUMS. (Referred in the main text Section 1.1)

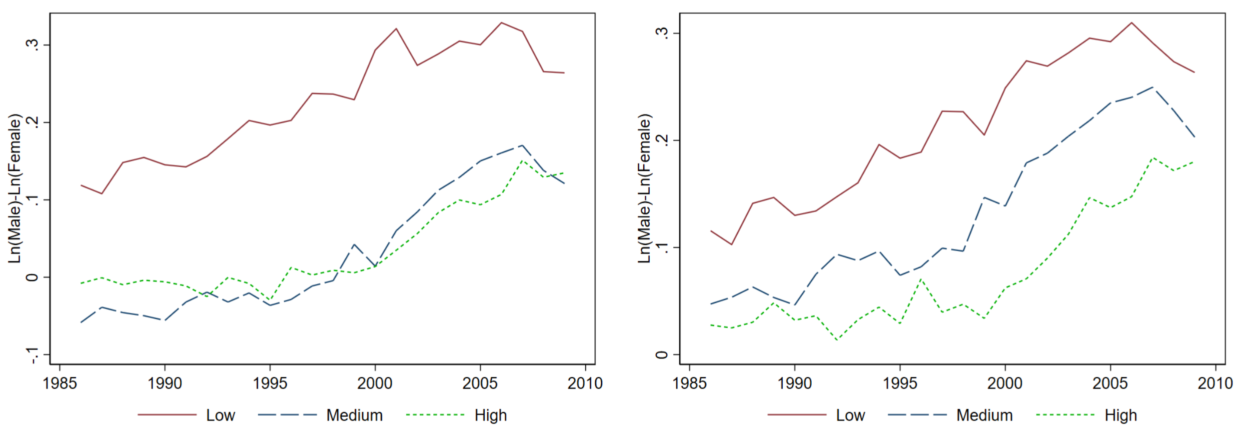



(a) Married women share

(b) FLFP by martial status

Figure A.2: Married women share and FLFP by marital status (aged 25-54)

Note: Marital status is determined by survey questionnaires. (a) The share of married women had declined but the majority of women aged 25-54 are married. (b) The FLFP of single women is higher than married women. (Referred in the main text section 2.1)




(a) Residual wage approach

(b) Gender dummy approach

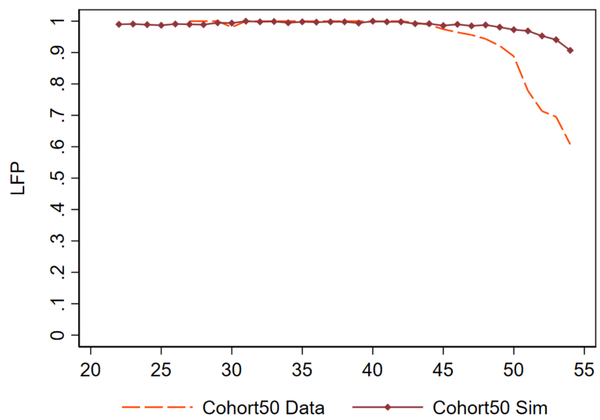
Figure A.3: Gender pay gap in alternative measurements.

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 2.3). (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 2.3)

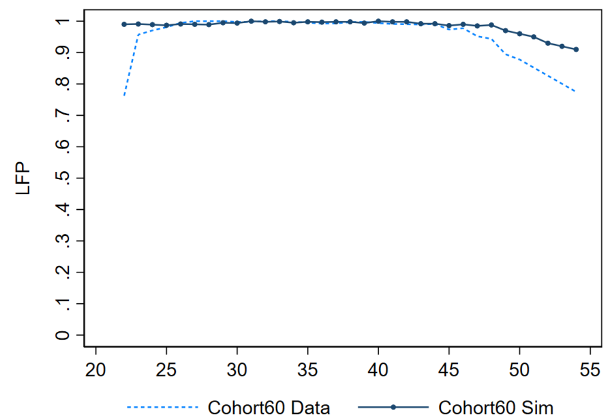
Table A.1: City-level Gender Wage Gap and FLFP

| Variables | (1)<br>ln(FLFP) | (2)<br>ln(FLFP) | (3)<br>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 |

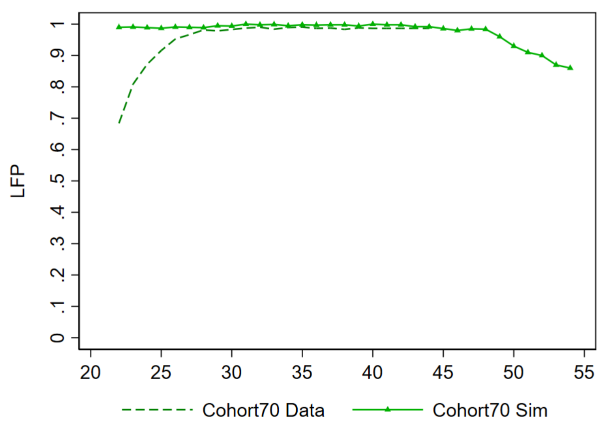
Note: Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The gender wage gap and the FLFP are negatively related at the city level. The second column is a preferred specification: 1% increase in the gender pay gap decreases FLFP by 0.2% which captures the scale in Figure 2.3. Controlling 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.4: Fit of high-educated women's FLFP (non-targeted moment).

Note: Income parameters are estimated from Equation (**??**). Other parameters come from medium-educated women of the same cohort. (Referred in the main text section 4.2)

Table A.2: Estimation of men's earning

| | | Cohort 50-59 | Cohort 60-69 | Cohort 70-79 |
|---|---|---|---|---|
| Low-educated | $\tilde{b_0}$ | 8.295 | 8.345 | 9.210 |
| | $\tilde{b_1}$ | 0.042 | 0.115 | 0.089 |
| | $\tilde{b_2}$ | 0.000 | -0.002 | -0.002 |
| Medium-educated | $\tilde{b_0}$ | 8.233 | 8.436 | 9.215 |
| | $\tilde{b_1}$ | 0.061 | 0.117 | 0.121 |
| | $\tilde{b_2}$ | 0.000 | -0.002 | -0.003 |
| High-educated | $\tilde{b_0}$ | 8.208 | 8.372 | 9.472 |
| | $\tilde{b_1}$ | 0.070 | 0.158 | 0.127 |
| | $\tilde{b_2}$ | 0.000 | -0.003 | -0.003 |

Note: Estimated with a Mincer-type equation (9). (Referred in main text section 4.1)

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 4.2)

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 4.3)

# 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 2.1)

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 Weight

The data is used *without* weight as no official weight is provided. However, one concern is the larger number of households surveyed during 2002-2009 may have a disproportionate impact on calculations of cohort-dependent FLFP and earnings. To address this, I weighed the data according to the officially reported urban population and examined the effect of weight in Figure B.1. Since this weight has little impact on the FLFP and earnings, I use the unweighted data.

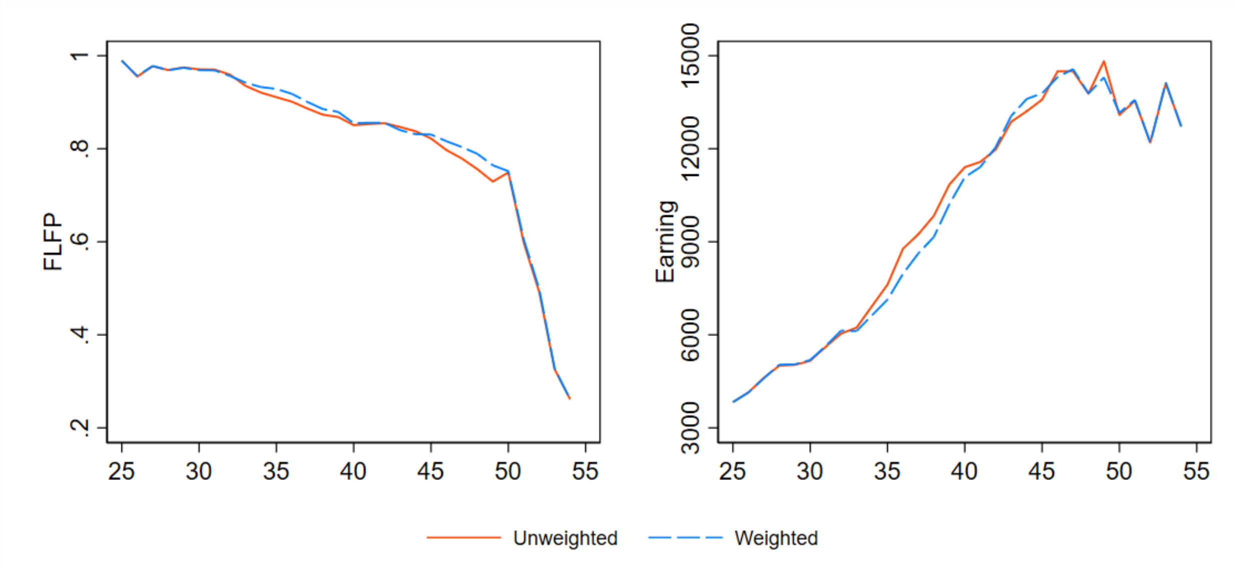



Figure B.1: Compare unweighted and weighted FLFP and earnings of the cohort 1960-1969.

Note: Cohort 1960-1969 is used to demonstrate the effect of weight as their lifecycle is fully covered by the data.

## B.3 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 FLFP 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.2 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 2.1)

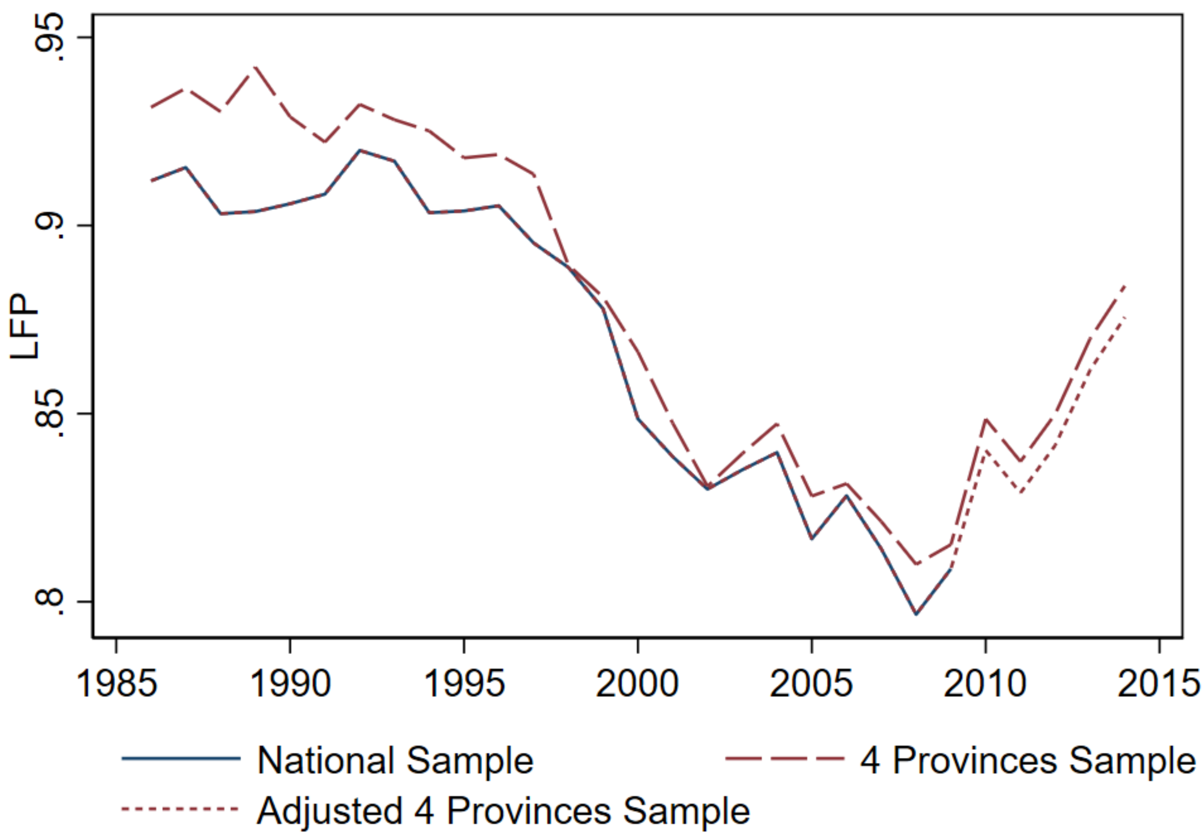


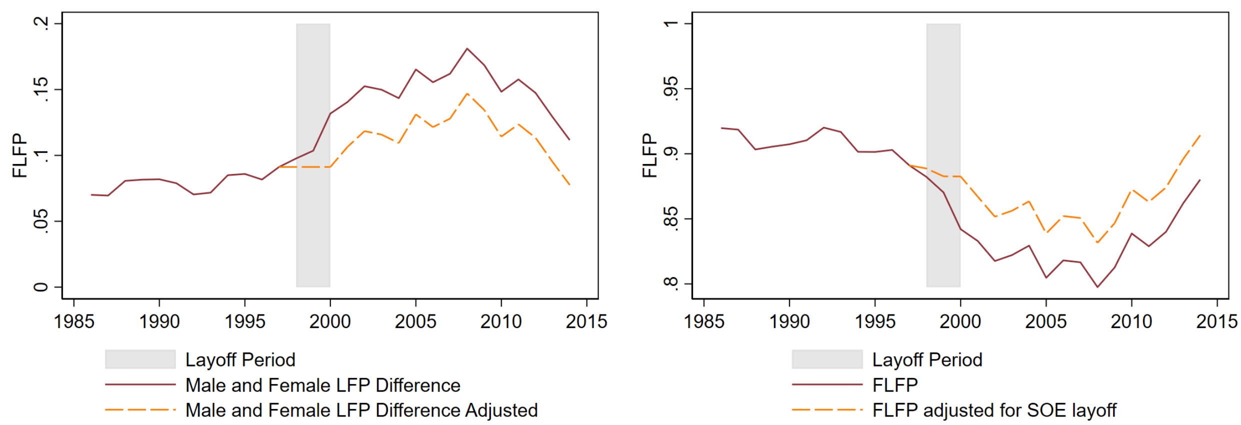

Figure B.2: Reconcile FLFP between Two UHS Data.

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.4 Adjust aggregate FLFP

The aggregate FLFP needs two adjustments. First, it needs to be adjusted for the massive layoff of SOE (state-owned enterprise) around 2000. Second, it needs to be adjusted for delayed retirement since 2010.

Figure B.3a shows the difference in LFP between men and women aged 25-54. There is a kink between 1998 and 2000, which coincides with the SOE layoff period 1998-2004 (Tian et al., 2022). I attribute all this gap to the SOE layoff and calculate FLFP as if there is no such shock (the orange dashed line in Figure B.3b). As all women were affected, it requires more assumption to attribute this adjustment to each age group in each cohort, which would only increase FLFP by 0.01-0.05% for certain groups. Therefore, I do not adjust the FLFP for each cohort.

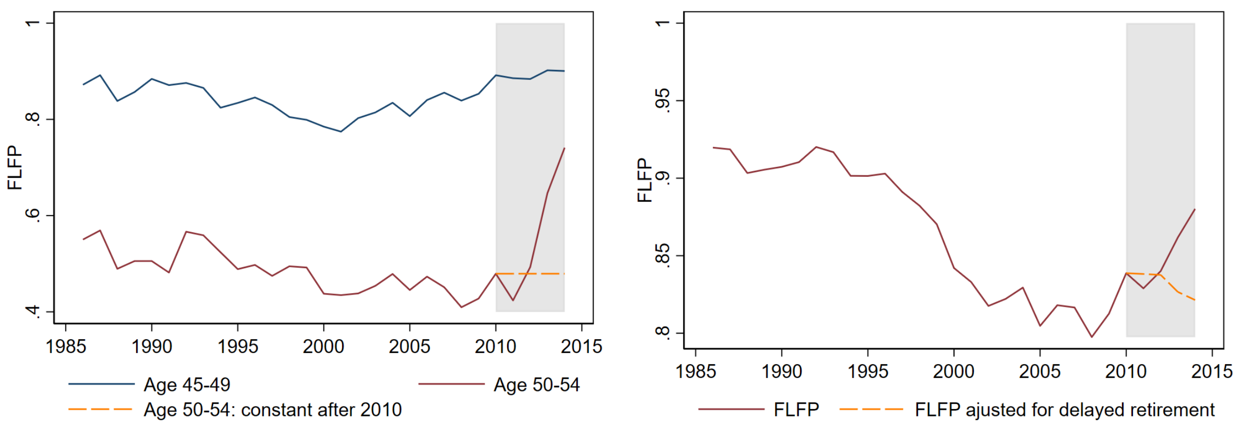



(a) (b)

Figure B.3: Adjust FLFP for the SOE layoff.

Note: The differences in LFP between 1998 and 2000 are all attributed to the SOE layoff.

The increased FLFP since 2010 is due to a higher FLFP among women aged 50-54. This delayed retirement itself is worth investigating but beyond the scope of this paper. Figure B.4a compares the FLFP of women aged 45-49 and 50-54 from 1986 to 2014. We can see that there is a significant surge in FLFP of women aged 50-54 since 2010. If I assume that the FLFP of women aged 50-54 is kept at the level of 2010, the increased FLFP since 2010 would reverse as the dashed line in Figure B.4b. (Referred in main text section 2.2)




(a) (b)

Figure B.4: Adjust FLFP for delayed retirement.

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

## B.5 Measure the Degree of Assortative Marriage

The separable Extreme Value (SEV) index is proposed by Chiappori et al. (2020). To construct the SEV index, consider the following scenario. Both masses of men and women are 1. The share of low-educated men is m and the share of low-educated women is n. r is the ratio of marriage between low-educated men and low-educated women, etc:

| Population Share | Low-educated women<br>n | High-educated women<br>1-n |
|---|---|---|
| Low-educated men<br>m | r | m-r |
| High-educated men<br>1-m | n-r | 1-n-m+r |

The SEV index is:

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

(Referred in main text section 2.5)

Table B.3: Degree of assortative marriage (SEV index) without age restriction.

| Cohort | Low-educated | Medium-educated | High-educated |
|---|---|---|---|
| 1950-1959 | 3.70 | 3.05 | 3.88 |
| 1960-1969 | 4.25 | 3.11 | 4.11 |
| 1970-1979 | 5.10 | 3.68 | 4.56 |
| $\Delta_{70-50}$ | 38% | 21% | 18% |

Note: SEV index is calculated based on all women without age restriction. Higher SEV values mean higher degrees of assortative marriage. (Referred in main text section 2.5)