University of California
Santa Barbara

Essays in Applied Economics

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Economics

by

Jahyeon Koo

Committee in charge:

Professor Shelly Lundberg, Chair
Professor Heather Royer
Professor Clément de Chaisemartin

June 2023

The Dissertation of Jahyeon Koo is approved.

Professor Heather Royer

Professor Clément de Chaisemartin

Professor Shelly Lundberg, Committee Chair

April 2023

Essays in Applied Economics

by

Jahyeon Koo

To my parents, Sanghoe and Insoon
who have always trusted and supported me

Acknowledgements

This dissertation would not have been possible without the help and support of many people. First, I would like to thank my advisor, Shelly Lundberg, who has consistently supported me in every step of my graduate study. I have benefited both academically and professionally from her thoughtful and instructive advice. I am grateful for her mentoring style — asking tough questions about my research while leaving me enough time to make further improvements. It would not have been possible to complete this dissertation without her support.

I would also like to thank Heather Royer and Clément de Chaisemartin, who have always provided helpful comments and suggestions for my research. Heather has helped me a lot whenever I was having a hard time. I am grateful for having them on my committee and learning from them throughout the last few years.

I want to thank Woongchan, Ken, and my cohorts. They always have supported me emotionally and academically. Moreover, I am in debt to the Applied microeconomics research group members, who contributed to my understanding of the empirical research process. I cannot forget Kieun Shim, my undergraduate advisor, who has encouraged me to pursue graduate studies and have provided practical advice during this period. Lastly, I am incredibly indebted to my parents, Sanghoe and Insoon. With their love and trust, I could finish this journey.

Curriculum Vitae

Jahyeon Koo

Education

2023 (expected)	Ph.D. in Economics, University of California, Santa Barbara
2018	M.A. in Economics, University of California, Santa Barbara
2017	M.A. in Economics, Seoul National University
2014	B.A. in International trade, Pusan National University
2013	Exchange student in Economics, University of Iowa

Research Interests

Demographic Economics, Labor Economics, Applied Econometrics

Working Paper

The Impact of Import Competition on College Choice: Evidence from Chinese Imports in Korea

Does a Pension Scheme Crowd Out Children? Evidence from Rural China

Revisiting Unilateral Divorce Law and Divorce Rate: Using Dynamic Effect Estimators with Heterogeneous Treatment Effects

Work in Progress

Regional Employment and Migration: Evidence from Chinese Imports in Korea

House Sales and Acquisition Tax in Korea

Honors and Awards

2022	Job Market Fellowship - University of California, Santa Barbara
2022	Research Quarter Fellowship -University of California, Santa Barbara
2013	Study Abroad Program School Scholarship, Pusan National University
2007, 2012	Merit-based School Scholarship, Pusan National University

Teaching Experience

Teaching Assistant - Principle of Microeconomics

Teaching Assistant - Principle of Macroeconomics

Teaching Assistant - Introduction to Econometrics

Abstract

Essays in Applied Economics

by

Jahyeon Koo

This dissertation consists of three essays. The first essay documents an empirical study of how the growth of imports from China changes post-secondary education decisions in Korea. As China has become "the world's factory", the availability of low-skilled jobs has become worse in countries that import products from China. I exploited cross-industry and cross-local labor market variations in import growth from China to investigate how Korean students' post-secondary education decisions were influenced by more intensive import competition from China. Using administrative statistics from high schools between 2000 and 2015, I found that in regions more affected by Chinese imports, more students pursued a college degree after finishing high school. Consistent with the enrollment pattern, I found that in Korea, a surge in Chinese imports caused huge job losses, especially for low-educated workers, which decreased the opportunity cost and increased the marginal benefit of a college education. However, there is heterogeneity in the enrollments based on gender and type of educational institution. In the more affected regions, male students tended to enroll in 4-year universities rather than 2-year colleges, while female students were more likely to choose 2-year colleges. The findings of this study suggest that negative social norms and smaller employment losses for working women led to a more limited increase in educational investment among women.

The second essay recounts an investigation into how parents change child-related investments in rural China after introducing a pension program that provides them with an alternative to adult children as a source of support. Exploiting regional variation in the timing of the New Rural Pension Scheme (NRPS) in China, I found that parents

who enrolled in the NRPS spent more on their children's education. Specifically, the increase in educational spending is observed for the parents of sons. In terms of the number of children, although the effect on the probability of giving birth to a baby is not statistically significant, I found that parents enrolled in NRPS were more likely to give birth to a boy. Considering the introduction of a pension scheme as lowering the price for future incomes, this implies that the income effect is more significant than the intertemporal substitution effect leading to higher investments in their children. Moreover, these results suggest that NRPS does not weaken son preference in China, although this program can provide old-age support for parents instead of sons who traditionally did this.

The third essay revisits literature estimating the effect of unilateral divorce law on the divorce rate with US panel data. This literature could not have a consensus on the impact of this law on the divorce rate because two-way fixed effects and event study regressions are not robust to heterogeneous treatment effects. In this case, estimates cannot be interpreted as a causal effect. By using alternative estimators to address this issue (e.g., de Chaisemartin and D'Haultfuille 2020), I found that the divorce rate rose after the adoption of this law, and this rise reversed afterward. The average of the dynamic effects is smaller than previous papers found.

Contents

Chapter 1

The Impact of Import Competition on College Choice: Evidence from Chinese Imports in Korea

1.1 Introduction

International trade drives countries to specialize in the production of the goods for which they have lower opportunity costs than other countries. By exporting a good for which it has a comparative advantage while importing another good, each trading country can increase its overall consumption. Consistent with the theory of comparative advantage, China has become the largest exporter of manufactured goods in the world based on abundant labor forces, resulting in considerable wage and job losses for low-skilled workers in countries that are China's partners such as the US (Autor et al. 2013, Autor et al. 2014). As the prospects of low-skilled workers become worse in countries that import products from China, educational attainment in these countries may change. This paper investigates how people in these countries adjust to rising Chinese imports in terms of education. In particular, this paper studies how high-school graduates in Korea responded to import competition from China by evaluating the changes in enrollment in Korean post-secondary institutions between 2000 and 2015.

The predicted effect of this change in trade patterns on educational attainment is ambiguous. On the one hand, as imports increase from China, more people choose to engage in education to obtain skills because the opportunity cost of obtaining skills decreases and the college premium for future earnings (difference in employment and earning between college graduates and high-school graduates) increases. On the other hand, fewer people may choose to attend college because their family incomes are negatively affected by job and wage loss. With imperfect credit markets and limits on borrowing, families experiencing negative income shocks from trade with China may not be able to finance tuition and living costs for college. Thus, students from low-income families will under-invest in their human capital.

In addition to the theoretical ambiguity, Korea is an interesting case to investigate, as it has several unique characteristics that distinguish it from other advanced countries. First, the economy of Korea depends heavily on trade. In 1990, foreign trade in the US consisted of 15.3% of the GDP, while in Korea, this sector accounted for 47.7% of GDP. Since China became a major supplier of manufacturing goods, Korea's trade dependency has become more severe. Trade value in 2019 reached 63.3% of Korea's GDP when the value in the US became 19.3%. During the same period, imports from China grew to 21.3% from 3.2% of Korea's total imports. Therefore, the emergence of China in international trade may have a more significant effect on Korea's specialization process than other countries due to their large dependence on trade. Second, Korean society has more traditional attitudes toward working women. According to the Integrated Values Survey in 1990 and 2000, Koreans were more likely to agree with the statement "When jobs are scarce, men have more right to a job than women" than many other advanced countries (Bertrand et al., 2020). Considering this negative social norm, China import shock may have caused Korean females to respond differently from Korean males, in comparison to females in other countries.

To determine how Korean students adjusted their college enrollment decisions when they graduated from a high school (most at 18-19 years old) in response to Chinese import shocks, I exploited regional variations in employment share of each industry to measure import exposure; if one region has a high employment share in industries importing largely from China, this region is highly affected by the China import shock. I compared the change in college and university enrollment between regions more and less exposed to this shock. The results show that a one standard deviation increase in local import penetration increased enrollment by 106.8 students per 100,000 residents between 2000 and 2015. This is a 10 percent increase in the enrollment rate relative to the average in 2000. Consistent with the increase in the enrollment rate, I found that in Korea, a surge in Chinese imports caused huge job losses especially for low-educated workers, which suggests that the opportunity cost and marginal benefit for college education became lower and higher respectively. As most importing industries were male-dominated, the import shock particularly reduced employments for low-educated men.

An increase in college enrollment is observed for both male and female students. A one standard deviation increase increased the enrollment by 55.1 male and 51.7 female students respectively. However, when investigating these effects by the type of college, I found that male and female students responded differently to the import competition: 60% of the increase in male students is at 4-year universities while the effect on female students is concentrated on 2-year college enrollments (77% of the increase in female enrollments). This is consistent with a larger reduction in employment for low-educated men, which lowered the opportunity cost and increased marginal benefit for male students more than for female students. Finally, I found that increases in female enrollment tend to be more limited in regions with more traditional gender norms for working women, suggesting an additional limitation on the expansion of higher education for women.

This paper is related to several previous works in the literature. First, this paper empirically investigates how trade affects post-secondary education. This study builds on previous theoretical analysis (Ferriere et al. 2021) as well as empirical evidence in the US exploiting variations of trade with China (Greenland and Lopresti 2016, Ferriere et al. 2021) and Mexico leveraging variations of tariff reductions due to NAFTA (Lee, 2021): Greenland and Lopresti (2016) and Ferriere et al. (2021) found that Chinese import competition increased high-school graduation rates and college enrollments while Lee (2021) found that tariff reductions due to NAFTA led to increase in community college enrollment rates in the US. This paper adds evidence from Korea. Moreover, compared to previous papers, this paper shows a greater effect in Korea, providing evidence that import shocks have a stronger impact on education in more trade-dependent economies.

Second, this article's findings contribute to the literature on how local economic conditions affect educational attainment. Previous research has found relationships between education and recession (Betts and McFarland, 1995; Barr and Turner, 2013 ; Foote and Grosz, 2020), housing booms (Lovenheim, 2011; Charles et al., 2018), local plant closures (Hubbard, 2018), local trade impacts (Lee, 2021). Most papers that investigated post-secondary education used college enrollments based on the location of colleges.[1] However, these estimates are likely to be biased because we do not know students' home regions when they decided to work or go to college. Specifically, enrollments in local colleges include students from different regions, which would cause a measurement error bias since they must be affected by economic shocks in the corresponding areas. Moreover, this excludes local students enrolling at colleges in different regions, which should be included to estimate the entire effect of local shocks. As almost half of students moved away from home for college in the US and Korea, these biases would be large. To correct these, I used the college enrollments based on the home regions when they decided to go to

[1] Recent papers such as Charles et al. (2018) and Ferriere et al. (2021) raised an issue of using the location of colleges and tried to solve this issue, for instance, by restricting sample to people living in their state of birth.

college: the location of the high school they graduated from. As Korean students are required to report to their high school whether they enroll at a college when they graduate, by using administrative statistics from local offices of education in Korea, it is possible to know where they live when they decided on post-secondary education. Although this measure only provides data on 18- and 19-year-old students, the estimates are likely to be close to the true overall effect of import penetrations on college enrollments in Korea, as most of new enrollments (about 80%) are new high school graduates.

Third, this paper is related to research on gender inequality in educational attainment. I found that an increase in imports from China did not contribute to reducing the gender gap. Although I found similar increases at all colleges among males and females, I found heterogeneous effects by gender and by type of institution; male students tended to choose 4-year universities, whereas females tended to enroll in 2-year colleges in response to import competition from China. This is consistent with the understanding that the industries most affected by imports from China were male-dominated, and thus the opportunity cost and marginal benefit of additional education for male students became smaller and larger than it became for female students. Another possible reason for the gender gap in enrollment is that Korean society has more negative attitudes toward working women, which may limit the perceived return to occupations requiring a 4-year college education. This paper found a significantly lower effect of import penetration on female enrollment in 2-year colleges in more traditional regions where people showed more negative attitudes toward working women.

The remainder of the paper is organized as follows. Section 2 explains backgrounds including education system in Korea, trade between Korea and China, and a theoretical background analyzing this problem. Section 3 introduces empirical strategy and data I used. Section 4 presents main results and additional analyses such as robustness check and heterogeneous effects to explain a possible mechanism for these results. Finally, section 5 summarizes findings and discusses further research questions related to this paper.

1.2 Background

1.2.1 Education in Korea

The Korean education is a single-track system, which operates on 6-3-3-4 basis, with six years of elementary school, three years of middle school, three years of high school, and four years at the undergraduate university level (two or three years at the junior college). There are nine years of compulsory education, including the six years of elementary school and three years of middle school.

Higher education in Korea is provided primarily by universities and colleges. The Korean higher education system is modeled after the United States with colleges (namely junior colleges or community colleges) awarding licenses, associate degrees, or diplomas while universities award bachelor's, master's, professional, and doctoral degrees. Specifically, junior colleges offer professional certifications and programs related to vocational education such as business administration, technology, engineering, agriculture, and nursing. After the program, most junior college graduates choose to enter the workforce as semi-skilled technicians and service workers. However, unlike the US, a very small number of the graduates transfer to a four-year university to further their studies since universities in Korea accept a small number of transfer students. On the other hand, universities provide theoretical education and research in various domains of knowledge and disciplines. After receiving a bachelor degree, they usually get white-collar or skilled technical jobs. Otherwise, they pursue a higher degree in graduate school.

Undergraduate tuition fees in Korea are relatively cheaper than the US. In 2005, the average tuition for 1 year in national and private universities were \$3,264 and \$6,024 respectively (Korea Higher Education Research Institute, 2013). However, student loans are common. In 2005, about 294,000 students used student loans, and that number more than doubled in the recession of 2008 to 635,000, which was 34% of college students (Corea Institute for New Society, 2009). For the last 5 years, the average rate has been about 14% (Ministry of Education, 2019).

Although many students needed to borrow money to finance tuition and living costs for college, it was not easy to get a student loan in the early 2000s. The Korean government tried to solve this problem since expensive tuition and low support for college students were one of the hottest political issues. Before 2005, students could not get a loan without a parental guarantee, which made students from low-income families impossible to pursue a college education. Even after the government-insured loan program was initiated in 2005 to solve educational inequality, commercial banks charged a high-interest rate, which led to another government reform.[2] The government implemented national scholarship and student loan programs with lower interest rates in 2010 and 2012 respectively to support students directly (Han and Kang, 2013).

Figure 1 shows the share of people holding college degrees between 2000 and 2015 in Korea and the US. Although the share in Korea was lower than that of the US before 2014, the growth was much faster than the US. Due to the higher growth in Korea during this period, the share of Korea surpassed that of the US after 2014. Figure 2 shows college enrollment rate by gender in Korea between 1991 and 2016. The growth in women's enrollment was even faster than in men's and after 2009 women's enrollment rate was higher than men's. After 2009, the enrollment rate of women was higher than that of men. One of the reasons for the extremely high growth rate was the large supply of post-secondary education in the 1990s. The government loosened restrictions on college establishments, which doubled admission capacities between 1990 and 2000.

[2]In 2009, the interest rate for a college student loan was about 7% while the average interest rate for other loans was less than 5% (Korea Federation of Banks, 2013).

1.2.2 Trade: Korea and China

There are many factors that have contributed to China becoming the largest exporter of goods in the world since 2009. Economic reforms in the 1980s and 1990s (Naughton, 2007; Hsieh and Ossa, 2016; Zhu, 2012), rural to urban migration flows of more than 150 million workers (Li et al., 2012), and massive capital accumulation (Brandt et al., 2012) have led manufacturing to expand at a breathtaking pace. In addition to internal factors, China's participation in the World Trade Organization (WTO) in 2001 increased their exports to the world (Naughton, 2007).

Like the share of China's exports to the world, the volume of imports to Korea from China has become much larger during this period. Since the two countries established a formal diplomatic relationship in 1987, imports increased from 1,705 million dollars in 1990 to 90,082 million dollars in 2015 (See Figure 2). The proportion of imports from China has grown significantly, to 20.7% in 2015 from 3.2% of total imports in 1990. Now, Korean imports from China are larger than those from any other countries. As China and Korea signed a free trade agreement in 2014, the amount of trade between these two countries is expected to continue to increase.

As in other countries, manufacturing employment has decreased as manufactured imports from China have increased. Figure 3 shows the growth of employment in manufacturing and non-manufacturing between 1989 and 2017 compared to employment in 1988. When non-manufacturing employment increased by almost 60%, manufacturing employment remained similar to 1988. This is similar to the trend in the US between 1990 and 2011, when the US increased imports from China significantly (Acemoglu et al., 2016).

In contrast to the similar patterns of Chinese imports for many countries, Korea has unique patterns. For instance, the Korean economy depended more on trade than other countries. Figure 4 shows the ratio of imports and exports to GDP in Korea and the US. While imports and exports comprised almost half of Korea's GDP, they made up only

15% of the US GDP. When both countries imported more and more from China, the Korean economy became more dependent on trade while the US economy's dependence on it remained stable. Figure 1 directly measures the ratio of import value from China to GDP in Korea and the US between 2000 and 2015. Similar to the trade dependence trends, Chinese imports became a larger part of Korean GDP, while its ratio to the US GDP was unchanged. These differences could be why the share of college graduates in Korea grew much faster than in the US: Due to the high dependence on trade, a higher proportion of students in Korea would be affected by Chinese import shock.

1.2.3 Theoretical Background

I used the model of Lochner and Monge-Naranjo (2012) to summarize the effect of import competition on post-secondary education. The effect depends on whether or not students can borrow money freely in financial markets to finance costs for additional education. If there is no borrowing constraint, educational attainment depends on student's ability, future incomes, and the opportunity cost of schooling, which depends on the wage they could earn if they work instead of schooling. However, if there is a borrowing constraint, educational attainment depends on their family wealth in addition to these.

To predict the effect of the rise in imports on education, consider two-period-lived individuals who invest in education in the first period and work in the second, and maximize their utility,

$$\max_{h,d} u(c_1) + \beta u(c_2)$$

$$\text{subject to } c_1 = W + w_1(1-h) - \tau h + d,$$

$$c_2 = w_2 a f(h) - Rd.$$

where c_t is consumption in period $t = 1, 2$, $\beta > 0$ is a discount factor, and $u(\cdot)$ is strictly increasing and concave. Each person is endowed with financial assets $W \geq 0$ such as family transfers, ability $a > 0$, and a unit of time. During the period 1, individuals make human capital investments h at tuition costs $\tau > 0$ to increase labor earnings $y = w_2 a f(h)$ in the second period.[3] Instead of schooling, they can work for w_1 in the first period.[4] Finally, they can borrow d at a gross interest rate $R > 1$ to finance τ. In a perfect credit market, there is no borrowing constraint, but in the imperfect market, there is the upper limit of borrowing, $d \leq \bar{d}$, which leads to a different theoretical prediction from the perfect market.

How does an individual optimize their utility in this problem? First, in the perfect market, students maximize their utility by investing in human capital until the marginal return on human capital is equal to the return on savings R, $MR(h^{p*}) \equiv \frac{w_2 a f'(h^{p*})}{w_1+\tau} = R$. As they can borrow money as much as they want, the optimal level of human capital h^{p*} does not depend on W. In contrast, if there is a limit for borrowing $d \leq \bar{d}$, there must be constrained students whose $W < W_{min}(a)$. Although these students have high abilities, because of low wealth W, they cannot cover tuition for their optimal level of human capital h^{p*}, $W + w_1(1-h^{p*}) + \bar{d} < \tau h^{p*}$. Therefore, they borrow money as much as possible $d^* = \bar{d}$ and their optimal level is $h^{i*} < h^{p*}$ such that $W + w_1(1-h^{i*}) + \bar{d} = \tau h^{i*}$. In this case, $MR(h^{i*}) \equiv \frac{w_2 a f'(h^{i*})}{w_1+\tau} > R$.

What prediction does this framework suggest about the effect of the import shock on educational attainment? Increased imports from China can be expected to decrease wages and available jobs for low-skilled people (Autor et al. 2013, Autor et al. 2014). This can affect h through three channels. First, a surge in Chinese imports lowers wages for students w_1, which leads them to invest more in h since the opportunity cost for additional h becomes lower (opportunity cost channel). Moreover, students may think that intensive

[3] $f(\cdot)$ is positive, strictly increasing, and concave.

[4] In Korea, military service could be another option for male students after graduating from high school. However, according to Military Manpower Administration, about 4.5% of enlisted soldiers were 19 years old (1998-2000). It implies that a small number of high school graduates choose to serve in the military service.

Chinese import competition lowers wages for low-skilled prime-age workers $w_2af(h^{low})$ persistently, which rises college premium for earnings $\frac{w_2af(h^{high})-w_2af(h^{low})}{h^{high}-h^{low}} \approx w_2af'(h^*)$ (college premium channel). Thus, these two channels predict higher educational attainment. However, the increase in import competition can lower family transfers, W, which lowers the college enrollments because more students cannot afford the tuition τ (wealth effect channel). Therefore, the theoretical analysis does not provide a clear answer for whether China import shock leads more or less enrollments at college.

1.3 Empirical Strategy and Data

1.3.1 Empirical Strategy

Basically, I compared the change in enrollments at any college between regions more and less exposed to China import shock. Specifically, I followed the empirical specification Autor et al. (2013) used, which is

$$\Delta(\frac{y_{ipt}}{pop_{ipt}}) = \beta_1 \Delta IP_{it}^{ck} + \mathbf{X}_{it}'\beta_2 + \gamma_{pt} + \epsilon_{it} \tag{1.1}$$

where the outcome is the change in enrollments of high-school graduates (most at 18 and 19-year-olds) per 100,000 people in a statistical metropolitan area[5] (SMA) i of province[6] p in period t (that is, between year t and year $t+5$).[7] In $\mathbf{X}_{it}$ of the preferred speci-

[5] SMAs in Korea are similar to those in the US, defined by commuting rates between cities and population. There are 50 SMAs with these criteria, and 95% people in 2000 lived in these SMAs (Kim et al., 2007).

[6] A province is a geographical division that includes SMAs.

[7] I used the total population in the denominator to make the estimates comparable with Lee (2021). When I used instead high school students, the results are similar to the main result.

fication, I included period-specific lagged manufacturing employment share,[8] the share of females, and the share of people aged between 15 and 29 to control for time-varying demographic characteristics in each SMAs. γ_{pt} represents fixed effects of province-specific periods to account for time-varying unobservables common within each province. ΔIP^{ck}_{it} is the measure of the local market shock, which is the average change in China import penetration in SMA's industries, weighted by each manufacturing industry j's share of SMA i in the initial year of period t, $\frac{L_{ijt}}{L_{it}}$:

$$\Delta IP^{ck}_{it} \equiv \sum_j \frac{L_{ijt}}{L_{it}} \cdot \Delta IP^{ck}_{jt}. \tag{1.2}$$

$\Delta IP^{ck}_{jt} \equiv \frac{\Delta M^{ck}_{jt}}{L_{jt}}$ is the growth of the imports per one worker in manufacturing industry j from China in Korea in period t. This is measured by the change of import values (in 1995 US dollar). The empirical strategy exploits cross-industry and cross-local labor market variation in import competition by import growth from China.

However, ΔIP^{ck}_{it} is likely to be correlated to unobserved local economic and demographic conditions. For instance, if the education level of Korea is high, this would lead to more imports and lower employment share of low-skilled goods. That is, due to reverse causality, β_1 could be biased. To deal with possible biases, I constructed the local import penetration ΔIP^{co}_{it}, using the growth of Chinese imports in other high-income countries. That is,

$$\Delta IP^{co}_{it} \equiv \sum_j \frac{L_{ijt-1}}{L_{it-1}} \cdot \Delta IP^{co}_{jt} \tag{1.3}$$

where $\Delta IP^{co}_{jt} \equiv \frac{\Delta M^{co}_{jt}}{L_{jt-1}}$.[9] Then, I used ΔIP^{co}_{it} as an instrumental variable (IV) for ΔIP^{ck}_{it}.

[8] Borusyak et al. (2021a) recommended to include the lagged employment share or the period-specific share when using industry shocks as an exogenous variation to prevent from leveraging non-random variation, $\sum_j \frac{L_{ijt-1}}{L_{it-1}}$ in the instrumental variable later.

[9] If L_t were used in ΔIP^{co}_{it}, this would lead to high correlation between ΔIP^{ck}_{it} and ΔIP^{co}_{it} by using

The idea of this IV is that the growth in Korea and other countries reflects only Chinese productivity shocks that are not correlated to unobserved characteristics in Korea. These productivity shocks, for instance, include the rising competitiveness of Chinese manufacturers and China's lowering of trade barriers due to China's economic reforms and WTO participation. Therefore, by using the IV, it eliminates the biases correlated to the unobserved conditions in Korea.

Thus, the empirical strategy is a shift-share design, using the import growth in the other countries as the source of exogenous variations. Although exposure shares $\frac{L_{ijt-1}}{L_{it-1}}$ would be endogenous, it is still possible to identify a causal effect if shocks ΔIP^{co}_{jt} are exogenous (Borusyak et al., 2021a).[10] The identifying assumptions for β_1 to be interpreted as a causal effect are a) ΔIP^{co}_{jt} is not correlated to unobserved industry characteristics of Korea (exogeneity)[11] , b) ΔIP^{co}_{it} is relevant to ΔIP^{ck}_{it} (non-zero correlation), and c) ΔIP^{co}_{it} affects the outcomes only through ΔIP^{ck}_{it} (exclusion restriction).

First, as this strategy depends on the exogeneity of import growth in the other countries across industries ΔIP^{co}_{jt} (Borusyak et al., 2021a), I assessed the exogeneity of ΔIP^{co}_{jt} by testing whether the import growth is correlated to pre-determined industry characteristics. For this, I used Mining and Manufacturing Survey in Korea that surveys char-

the same weight, $\frac{L_{ijt}}{L_{it}}$ not from between ΔM^{ck}_{jt} and ΔM^{co}_{jt}.

[10] To show this formally, consider a simple regression $y_i = \beta_1 x_i + \epsilon_i$ with $z_i = \sum_j s_{ij} g_j$ as an IV for x_i using g_j as an exogenous variation. Then,

$$\beta_1 = \frac{\sum_i z_i y_i}{\sum_i z_i x_i} = \frac{\sum_i \sum_j s_{ij} g_j y_i}{\sum_i \sum_j s_{ij} g_j x_i} = \frac{\sum_j g_j \sum_i s_{ij} y_i}{\sum_j g_j \sum_i s_{ij} x_i} = \frac{\sum_j s_j g_j \bar{y}_j}{\sum_j s_j g_j \bar{x}_j},$$

where $s_j = \frac{\sum_i s_{ij}}{N}$ are weights capturing the average importance of shock g_j, and $\bar{v}_j = \frac{\sum_j s_{ij} v_i}{\sum_i s_{ij}}$ is an exposure-weighted average of v_i. Thus, β_1 is equivalent to estimating the second-stage coefficient from a s_j-weighted shock-level IV regression that uses the shocks g_j as the instrument (See Proposition 1 in Borusyak et al., 2021a).

[11] In addition, Section 4.2.4 checks the balance across regions by investigating the correlation between ΔIP^{co}_{it} and pre-determined local characteristics. I fail to find any correlation.

acteristics of establishments in mining and manufacturing industries. In particular, I used characteristics from the survey in 1999, 2004, and 2009 to test exogeneity of ΔIP^{co}_{jt} for 2000-2005, 2005-2010, and 2010-2015.[12] For instance, as the characteristics in 1999 are pre-determined with respect to ΔIP^{co}_{jt} for 2000-2005, there should be no correlation between them.

Table A4 shows estimates of regression of ΔIP^{co}_{jt} on a set of pre-determined industry characteristics. I failed to find statistical significance in any coefficients of the characteristics. When I tested the joint hypothesis that the coefficients of all industry characteristics are equal to 0, the p-value is 0.29. Thus, I failed to find systematic evidence that ΔIP^{co}_{jt} is not exogenous in the instrumental variable ΔIP^{co}_{it}.

Regarding the second assumption (non-zero correlation), Table A2 provides evidence supporting non-zero correlation between the instrument and treatment variable. Regardless of clustering at the regional and industry level, the effective first stage F-statistic of Olea and Pflueger (2013) is larger than 10, which is the threshold to detect weak instrument (Andrews et al., 2019).

Finally, the exclusion restriction is related to the choice of countries for the instrument. If Korea economy is interdependent with other high-income countries, it may lead a direct effect on the outcomes, which violates the exclusion restriction. As there is no direct test for this assumption, I chose high-income countries that are frequently used in previous papers to study the effect of rise of import competition from China (Autor et al. 2013, Dauth et al., 2014, Balsvik et al. 2015, Helm, 2019). Those are Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, Canada, Singapore, United Kingdom, Norway, and Sweden. And then, I did experiments by dropping several countries in the instrument and compare with the main results. If the estimates after dropping some countries are so different from the main estimates using 12 countries, this could be evidence that the exclusion restriction is violated. First, I excluded Australia, Germany,

[12] I also failed to find statistical significance when I use pre-period changes in industry characteristics.

and Japan when constructing the instrument. As these countries were in the 10 most trading countries of Korea, and thus one can imagine that they are so interdependent with Korea. Panel B in Table A5 shows that excluding these countries does not change the main results qualitatively although the first stage F-statistic becomes lower. In Panel C, I excluded Asian countries, Japan and Singapore, and the results are similar to the main results with the low F-statistic again.

1.3.2 Data

Previous papers (Charles et al., 2018; Lee 2021) used college enrollment data based on colleges' locations. Instead, I chose to base my findings on college enrollments on students' high school location since this is the location they lived in when they decided to go to college. It assumes that high school graduates are less likely to consider working in other regions when they make a college decision. Considering the propensity to migrate rises with education (Molloy et al. 2011), this assumption is reasonable. During the sample period, 75% of high school graduates aged between 19 and 22 worked in regions in which they lived 5 years ago (Korean population and housing census).

College enrollments on the location of high-schools are available in Regional Educational Statistic Datasets by School Types in Korea; each year, all schools report their student data to their respective regional offices of education. In particular, high schools are required to report whether their graduating students enroll in college and the type of college. Therefore, the dataset held by the regional offices of education would primarily reflect information on 18- or 19-year-old students, which is the age of high school graduates in Korea. These data are aggregated at the regional level and made available on the Korean Education Statistics Service website. From this dataset, I used the regional-level data on the number of high school graduates who enrolled at 2-year junior colleges and 4-year universities in each year by gender. I used data from the years 2000, 2005, 2010, and 2015. Using this data has both benefits and drawbacks relative to the previous papers available in the literature.

A main benefit of the college enrollments based on the high school location, is that it is possible to identify where students were living when they decided to go to college. In contrast, the location is not available if I use the enrollments based on the location of colleges. To be specific about why this is important, note that the region-i-colleges' enrollments do not include region-i-students' enrollments at region-i'-colleges but include region-i'-students' enrollments at colleges in region i. Region-i-students' enrollments must be affected by the import competition in region i while region-i'-students' enrollments must be affected by the competition in region i' through wages for low-skilled workers and their wealth in corresponding regions. Therefore, to the extent that the region-i-colleges' enrollments exclude the former enrollments and include the latter enrollments, estimates using this outcome must be biased.[13] According to Eagan (2016), 48.4% of US college freshmen in 1990 enrolled in colleges over 100 miles away from their permanent home. This number has remained relatively stable over time and was 50% in 2015. This is similar in Korea (40.5%).[14] Thus, as the bias would be large by using enrollments based on the location of colleges, it is beneficial to use enrollments based on high school they graduated from.

The disadvantage of this data, however, is that since the enrollment information is collected when students graduate from high school, respondents are mostly 18- or 19-year-old. To understand the full effect, it would be necessary to include enrollments by students of all ages. For instance, 24-year-old males who lost their jobs may enroll in college to improve their skills, but this would not be captured by the outcomes in this paper. Therefore, this paper only estimated the full effect of imports from China

[13] Charles et al. (2018) and Ferriere et al. (2021) excluded region-i'-students' enrollments at region-i-colleges by restricting sample to people living in their state of birth. However, this strategy cannot still provide the entire effect of local economic shocks since their sample does not include region-i-students' enrollments at region-i'-colleges.

[14] University News Network, 2014. `https://news.unn.net/news/articleView.html?idxno=139682`

on total college enrollments among new high school graduates. However, as 40.5% of Korean students enrolled at colleges in different regions and new high school graduates make up 77.4% of all freshmen enrollments in Korea,[15] the benefit of using this data is comparatively more demonstrative of the overall majority.

For the treatment ΔIP_{it}^{ck} and instrumental variables ΔIP_{it}^{co}, I used UNcomtrade and Census on Establishments. UNcomtrade provides the dollar amount of each country's imports by industry level and partner country. I extracted the imports from China of Korea and twelve other countries in 2000, 2005, 2010, and 2015. The Census on Establishments surveyed employments in almost all Korean establishments. This is panel data at the establishment-level and is available by year. I used the total employments of each establishment in 1995, 2000, 2005, and 2010. Because UNcomtrade and the Census on Establishments use different industry codes (Harmonized System (HS) code and Korea Standard Industry code (KSIC)), I used the correspondence table between the two codes provided by Statistics Korea. Multiple 6-digit HS codes corresponds to one 5-digit KSIC code. In this analysis, the number of manufacturing industries by KSIC code I used is 304.

Finally, I calculated regional demographic variables for control variables in (1) and outcome variables for additional analyses from Korean population and housing census 2 percent sample between 2000 and 2015, which is similar to Integrated Public Use Microdata Series (IPUMS) census data in the US. Regional population sizes each year are obtained Korean Statistical Information Service.

1.3.3 Descriptive Statistics

Before showing the main results, I briefly summarized the descriptive statistics in Table 1. On average, per 100,000 people, 990 students enrolled at any college, and almost 70% of them enrolled at a 4-year university. This pattern was observed in both male

[15] Korean Educational Development Institute, 2015.

and female students (Table 1-a). As shown in Figure 2, the gender gap between male and female students became smaller during 2000-2015, and in 2015, female enrollments overtook male enrollments at any college (Table A1-a).

Table 1-b shows descriptive statistics about regional characteristics in the initial year of each period. Using Korean imports, the increase in manufacturing imports from China ΔIP^{ck}_{it} is 970 dollars per worker, and if I use imports from other countries, ΔIP^{co}_{it} becomes 5,110 dollars. Table A1-b shows this by year; although imports from China became larger, the growth rate slowed between 2010 and 2015 as the growth of Chinese imports in Korea and other countries diminished. This pattern is also observed in Figure 5, which shows the regional variation in import penetration. As Korea imported more manufacturing goods, the share of people in manufacturing fell, while that of the employed increased. This implies that Korea has begun to specialize in non-manufacturing industries (Table A1-b).

Finally, Table A2 shows the manufacturing characteristics in 2000. On average, each establishment employed 21.7 workers and 69% were male. The last three rows in Table A2 show the relationship between the number of imports and the characteristics of employment. Industries that are more affected by an increase in Chinese imports tend to be more labor intensive and male dominated. This demonstrates a higher effect of import competition on males.

1.4 Results

All regression models here are weighted by the population size of SMAs, and I mainly reported standard errors clustered at the regional level. I also reported standard errors at the industry level in Table A3.[16] The result is not different from the main result.

[16] To calculate standard errors clustered at the industry level, I ran shock-level (industry-level) regression by using `ssaggregate` in Stata (Borusyak et al., 2021a).

1.4.1 Main Results

Figure 6-a and 6-b show scatter-plots between ΔIP_{it}^{co} and ΔIP_{it}^{ck}, and between ΔIP_{it}^{co} and the outcome after removing the effects of lagged manufacturing shares and period fixed effects. As expected, the local measures of Chinese imports in other countries are strongly correlated to the measures of the imports in Korea (Figure 6-a). Figure 6-b supports that the opportunity cost and college premium channels were more powerful than the wealth effect channel. As regions were more affected by imports from China, more high-school graduates enrolled at colleges. The ratio of two coefficients is an IV estimate 197($= \frac{45.6}{0.23}$) (first column in Panel A of Table A3), which implies that a one standard deviation increase in ΔIP_{it}^{ck} led to higher enrollments at any colleges by 197 students per 100,000 residents.

Table 2 shows the estimates of the local import penetration caused by China on post-secondary institution enrollment by using (1). In the first column, I measured the effect on enrollment at all post-secondary institutions in Korea. The estimate represents the change in the enrollment rate due to the difference in the import penetrations among SMAs. The estimate of 106.8 implies that a one standard deviation increase in the local import penetration increased enrollment by 106.8 students per 100,000 residents between 2000 and 2015. Alternatively, this increase is equivalent to a 9.6 percent increase relative to the average enrollment of 1110.5 students per 100,000 residents in 2000. To gauge this estimate, I ranked all SMAs by the change in local import penetrations from 2000 to 2015. The 75th percentile SMA experienced a 2.05 standard deviation greater increase in average local penetrations than the 25th percentile SMA. Therefore, the 75th percentile SMA is expected to have 218.94 more students in post-secondary institutions per 100,000 residents in 2015 relative to an SMA in the 25th percentile holding the 2000 enrollment levels constant. Alternatively, this increase in the enrollment rate corresponds to 19.7 percent of the 2000 enrollment levels.

I divided the total effect into the effect on 2-year college (junior college) and 4-year university enrollment. From the effect on any college enrollment (106.8 students per 100,000 residents between 2000 and 2015), a one standard deviation increase in the import penetration increased 2-year college enrollment by 61.2 students while increasing 4-year university enrollment by 45.7 students per 100,000 residents. However, I failed to find statistical significance when clustering at the SMA level. Again, comparing the 75th percentile with the 25th percentile SMA in import penetration, the 75th percentile SMA is expected to have 125.5 more students (93.7 more students) in 2-year college (in a 4-year university) per 100,000 residents in 2015 relative to an SMA on the 25th percentile, which is 36 percent higher enrollment (12.3 percent higher enrollment) in the 75th percentile SMA.

How large are these effects compared to other countries? Ferriere et al. 2021 used the Panel Study of Income Dynamics (PSID) data to estimate the effect of the China shock in the US. Comparing the 75th percentile commuting zones (CZs) with the 25th CZs, the probability of a high school graduate going to college is higher by 6.8 percentage points between 1990 and 2007.[17] In our case, the 75th percentile SMA is expected to have 12.1 percentage points higher during 2000-2015.[18] Another comparable estimates with us is Lee (2021)'s estimates. Among her estimates, the most comparable one is their estimate on 2-year college enrollment of students under age 20. Again, comparing the 75th percentile with 25th percentile CZs in the local tariffs of the North American Free Trade Agreement (NAFTA), Lee (2021)'s estimate (Table 4) shows that the 75th percentile CZs is expected to have 52.89 more community college students per 100,000 residents, which is about 42.1% of the estimate in Table 2. Compared with these two papers studying education in the US, the import competition effect on US students is approximately half of the effect in Korea. This supports that import competition affects

[17] 0.045 (in the column 1 of their table 4) $\times$ 1.51 (ΔIP_{it}^{cu} from Autor et al. (2013)) = 0.068.

[18] To make our estimate comparable with Ferriere et al. 2021, I used the ratio of college enrollments to high-school students as an outcome. 0.059 (the standard error is 0.015) $\times$ 2.05 = 0.121.

students in a trade-dependent economy more significantly.

The next question to answer is as follows: Did male and female students in Korea adjust differently to the China import shock? The growth rate of college enrollment among females was higher than that among males from the 1990s. Moreover, after 2009, the college enrollment rate among women remained higher than that among men (Figure 2). During the same period, the growth rate of imports from China was substantial (Figure 2). The correlation between the gender gap and imports may imply that trade with China has somehow contributed to reducing Korea's gender gap in educational attainment. In contrast, the industries most affected by imports were male-dominated; one may expect to widen the gender gap in college enrollment. In addition to whether male or female students reacted more strongly to the China import shock, one may expect heterogeneous effects by gender across institution types.

Panel B in Table 2 presents the import penetration effects by gender and by institution type. While a one standard deviation increase in the import penetration increased male enrollment by 55.1 male students, it led to an increase in female enrollment by 51.7 female students per 100,000 residents. The import competition effects on education are not heterogeneous by gender. However, while the effect on male students was concentrated in 4-year universities, most female students enrolled at 2-year colleges: Among students who enrolled at any college, 61% of males enrolled at 4-year universities while only 23% of females enrolled at this institution due to the Chinese import shock.[19] This implies that the import shock widened the gender gap in educational attainment by leading males to obtain a higher degree.

Why did more Korean students enroll in colleges? Two factors may have contributed

[19]To test whether the difference in the ratios is equal to 0 or not, I estimated the effects on any college and 4-year university enrollments for male and female students with the reduced form regressions, using seemingly unrelated regressions. And then, I calculated the standard error of the difference in the ratios based on the delta method. The difference is significantly different from 0 at 10% level (P-value of the difference: 0.068).

to this: 1) import competition from China may have lowered the opportunity cost of additional education w_1, and 2) the import competition may have increased college premium $w_2 a f'(h^*)$ students considered. In the next section, I investigate whether these factors contributed to more college enrollments in Korea.

1.4.2 Additional Analyses

Suggestive Evidence on Underlying Mechanisms

- **Opportunity Costs for Additional Education and College Premium:**

In this subsection, I examined how the penetration from Chinese imports affects local employment rates. As the local economy is an important factor when deciding between working and attending college after high school graduation, I estimated the effect of import competition on local employment rates to determine whether the change in local labor outcomes is consistent with the enrollment patterns that I found.

To do this, I used the Census on Establishments and the Korean Population and Housing Census 2 percent sample. As the Census on Establishments has employment records of all registered establishments and their industries, it is possible to determine which industries are more affected by a surge in imports from China.[20] However, as this census does not have any demographic information on workers except gender, it is impossible to determine if employments were differently affected across educational levels. To investigate this, I used the Korean Population and Housing Census 2 percent sample, especially males aged 25−34 and females aged 23−32. Individuals in these age ranges are likely to finish post-secondary education,[21] and thus I believe that labor market

[20]The average of employment rates in Census on Establishments is lower than the average in Korean population and housing census because Census on Establishments surveys only registered establishments, which is about 4 million establishments out of 7 million.

[21]According to Korean Educational Development Institute (2015), most females graduated from 4-year university at 23 while most males graduated at 25.

opportunities for these workers are considered as relevant by younger cohorts making education decisions. In addition, I restricted the sample to individuals who live in the same regions five years prior to mitigate concerns about endogenous migration. Finally, to the best of my knowledge, since there is no large regional data on wages in Korea for the sample period, I focused on employment rates rather than wages.

Table 3 shows the effects of local import penetration on employment rates in SMAs by gender and industry. Panel A of Table 3 shows that an increase of one standard deviation in local penetration reduced the employment rate by 2.9 percentage points, which amounted to 7.8% in 2000. Moreover, I found negative effects on both manufacturing and nonmanufacturing industries,[22] although the effect on the latter was insignificant. The significant effect on manufacturing employment is consistent with the fact that most imports from China were manufactured goods.

Table 4 shows the effects of local import penetration across genders and education levels based on individual-level data. An increase in import competition also reduced the employment rate among individuals who were likely to finish post-secondary education by 1 percentage points. People without a college degree were particularly affected by Chinese imports. An increase of one standard deviation in local penetration reduced the employment rate for people without a college degree by 1.9 percentage points. In contrast, I found a small and insignificant effect on employment rates for people with any college degree. These suggest that since employments for less-educated people became worse after a surge in imports from China, more Korean students enrolled in college due to the lower opportunity costs and higher marginal benefits of college.

Although it is impossible to investigate the effect of import competition on wages since regional wage data were unavailable in Korea during the sample period, import

[22] This is consistent with the findings of Asquith et al. (2019) in the US. They argued that the spill-over effect on non-manufacturing industries from manufacturing imports from China is evidence of negative aggregate demand effects.

competition would not have different effects on wages. For example, Ferriere et al. 2021 showed that an increase in Chinese imports reduced wages and employment in the US, with more profound effects on low-educated workers. Furthermore, using industry variation in Chinese imports into Korea, Ok et al. (2007) found that the wage gap between low- and high-skilled workers widened in more massively importing industries. Thus, based on the findings in this and previous papers, the increase in educational attainment during the China shock was likely to come through lower opportunity costs and higher marginal benefits for college as a surge in Chinese import worsened employment opportunities and wages for low-educated workers.

An analysis of these effects by gender revealed that import competition reduced the male employment rate by 2.7 percentage points, with the negative effects mostly concentrated in manufacturing industries (Table 3). This is unsurprising, given that manufacturing industries are male-dominated. Conversely, the effects on the rate of female employment were statistically indistinguishable from 0. In total, an increase of one standard deviation in the local penetration reduced the female employment rate by 0.2 percentage points, which accounted for 8% of the effect on the male employment rate ($= \frac{0.2}{2.7}$).

The second and third columns in Table 4 show the effects of local import penetration across genders and education levels. Among males who were likely to finish education, an increase of one standard deviation in local penetration reduced the employment rate by 1.2 percent. The effect was most pronounced on males without a college degree. Females were less affected by Chinese imports: an increase of one standard deviation increased the female employment rate by only 0.2 percentage points. I also failed to find significant effects on high-educated men and women due to large standard errors. These results suggest that more male students pursued higher education than female students because the negative impacts were concentrated on low-educated men, which led to lower opportunity cost of additional education and higher college premium for male students than for female students.

Table 4 could be evidence of why male students invested more in human capital than female students did, but it cannot explain why more female students enrolled at 2-year colleges. As import competition did not change female employment much in Korea, female students may have had little incentive to pursue post-secondary education. One possible scenario consistent with our results is that they were more motivated to study for vocational occupations, such as service jobs, that would provide them an income if and when their future husband's well-paying jobs disappeared.[23] However, it is impossible to determine this hypothesis with the current data.

- **Social Attitudes toward Working Women in Korea:**

I investigated how the effects of local import penetration differed based on social norms toward working women. There have been many attempts to ascertain how traditional gender stereotypes and norms influence gender gaps in education. For instance, because of conventional notions that men outperform women on standardized tests, especially in mathematics, women experience heightened anxiety during test taking that interferes with their test performance (Steele, 1997; Spencer et al., 1999). Social norms could affect motivations for education as well. If a society has a negative social norm in regard to working women, women would be less motivated for education (Bertrand et al., 2020), thus widening gender gaps in educational attainment.

Korean society has a more negative social attitude toward working women than other western countries. In addition to lower female labor force participation, Korean women cannot easily return to work after having a child compared to women from other devel-

[23]For example, Besedes et al. (2021) found that more women entered the labor force to offset lost incomes due to their male partners' exit from the labor force after the US granted China permanent normal trade relations status. Since more men in Korea lost their jobs after the increase in Chinese imports, female students may have been motivated to pursue more education for securing employment that could enable them to provide for their future households.

oped countries, which means that they suffer more considerable employment penalties due to having a child.[24] Moreover, college graduate women who earn more money than high school graduates are less likely to marry, while the opposite pattern has been observed in the US (Hwang, 2016).

Social norms/attitudes and child penalties are directly relevant to women's college premiums. If a society has negative social attitudes toward working women, the marginal benefits of additional education for women is because women who choose to work outside the home are subject to additional psychological costs (Bertrand et al., 2020). Similarly, in a society with considerable child penalties, the expected benefits of further education for women would be lower since women find it difficult to return to work after childbirth. Thus, college premiums female students perceived may have been lower than the premiums male students considered due to the traditional norms in Korea.

To examine whether social norms affected the responses of Korean women, I compared the effects of import penetration based on social norms within Korea. Following Hwang et al. (2019), I used regional sex ratios at birth (male-to-female ratios) to measure negative social attitudes toward working women in Korea. Hwang et al. (2019) used the sex ratio at birth for several reasons. In regions with higher sex ratios, women have lower educational attainment and are less likely to be employed and to hold leadership positions in governments, which is consistent with the social norms. Second, the ratio of male-to-female births captures the prevalence of son preference or, more broadly, traditional gender norms in a region. Individuals born in regions with high sex ratios at birth are more likely to disagree with the statement "dual-earner couples should equally divide housework" (p.6, Hwang et al., 2019). This statement measures the traditional norm for working women like "when jobs are scarce, men have more right to a job than women" in

[24]According to Korea Economic Research Institute (KERI), the employment rate of females with a child under 15 was 57% in 2019, which is much lower than the average of other developed countries, 72.2%. This is the third lowest number out of 40 developed countries (Figure A1).

Integrated Values Survey (IVS). Responses to the IVS statement have been widely used to measure egalitarian gender attitudes (e.g., Campana et al., 2018).

While Hwang et al. (2019) used sex ratios at birth across provinces in the early 1990s, I used the sex ratio across SMAs in 2000. Using a two-percent sample of the Korean population and housing census, I found patterns similar to those observed by Hwang et al. (2019). In 2000, women in SMAs with high ratios[25] were less likely to have a college degree and to work than those in SMAs with low ratios.[26] Moreover, in these regions, high-educated women were less likely to marry than those in SMAs with low ratios relative to high-school graduates.[27] These patterns are consistent with the predictions of Bertrand et al. (2020)'s theoretical model that the relative deficit that skilled women experience in the marriage market is larger in societies with more traditional gender attitudes, leading fewer women to pursue higher education. Thus, I believe that the sex ratio at birth in 2000 is a good measure of social norms that disadvantage working women.

Panel B of Table 5 shows coefficients of IV and the interaction between the IV and SMA with high sex ratios at birth (Panel A shows the coefficient of the IV to compare). I reported reduced-form estimates due to many-weak instruments problem. The effects of import penetration on female enrollment, especially in two-year colleges, were significantly weaker in more traditional regions than in less traditional ones. In the former, females did not enroll in two-year colleges in response to import competition. However, this was not observed in the case of male enrollment in four-year universities, where most male students enrolled. These patterns are consistent with my hypothesis that although women could obtain a college premium by pursuing a college education, women in more

[25]Those are regions with sex ratios higher than the median SMA ratio of 110.13 in 2000.

[26]While 19.14% and 38.61% women aged older than 23 in SMAs with high ratios had a college degree and worked in 2000, 21.53% and 42.37% of them had the degree and worked in SMAs with low ratios.

[27]Specifically, in SMAs with high ratios, among women aged 22−44 years (marrying age according to Shenhav, 2021), college graduates were 24.3 percent less likely to marry than high school graduates, while the corresponding difference in SMAs with low ratios was 23 percent.

traditional regions are less motivated to receive additional education due to the additional psychological cost they bear when working.

However, it is impossible to rule out the possibility that other unobserved factors related to traditional norms explain our estimates. For instance, people who adhere to such norms may disregard vocational jobs, which may have had null effects on two-year college enrollment for both male and female students. Thus, the estimates shown in Table 5 can be considered solid but not irrefutable evidence of the effects of the interaction between social norms and import competition.

Bias by Using Enrollments at Local Colleges

This subsection gauges the biases when using college enrollments based on the locations of colleges used in previous papers as an outcome (Charles et al., 2018; Lee, 2021). Enrollments at local colleges include students from different regions but exclude local students enrolling at colleges in different regions. Due to the former, enrollments at local colleges are an outcome variable with measurement errors. If these errors are correlated to our instrumental variable ΔIP^{co}_{it}, we do not know whether our estimates show the effect on college enrollments or the correlation between the errors and ΔIP^{co}_{it}.

To measure the significance of the biases from measurement errors, I used the Korean population and a housing census. Specifically, to compare estimates using enrollments at local colleges, I used local college enrollments of 18-19-year-old students who still live in the same regions they lived in 5 years ago. This outcome removes bias by excluding students who came from different regions for college within the last 5 years. Thus, the difference between these two estimates measures the bias from the measurement error. As local students' enrollments at local colleges do not include their enrollments at non-local colleges, the estimate must be the lower bound of the local import penetration effect.

Table A6 shows an estimate of the effects of import competition ΔIP^{co}_{it} using enrollments (local + non-local students) and local students' enrollments at local colleges.

Compared to local students' enrollments, the estimate of enrollments at local colleges is lower by 10.4% ($= \frac{27.66-30.87}{30.87}$). This implies that students who enrolled at local colleges from different regions, which is the measurement errors, are negatively correlated to the instrumental variable ΔIP^{co}_{it}. Moreover, since the estimate of local students' enrollments at local colleges is the lower bound, this implies that using all enrollments at local colleges underestimates the effect of import competition by at least 10.4%.

Note that the measurement error bias could be sensitive to empirical specifications. When I controlled for period fixed effects (Column 3 and 4 in Table A6) instead of province-specific period fixed effects (Column 1 and 2 in Table A6), the estimate of all enrollments at local colleges is larger than the estimate using local students' enrollments at local colleges by 22.6% ($= \frac{70.37-57.41}{57.41}$). This implies a positive correlation between the measurement error and instrumental variable. Thus, the sign and magnitude of biases from the measurement errors are sensitive to empirical specifications, which makes it hard to infer the actual effect of import competition on college enrollments. This would be also the reason why previous papers (for instance, Lee, 2021) could not find a significant effect on 4-year university enrollments in the US.

Robustness Check

- **Confounding factors**

This subsection investigates whether or not confounding factors drove our results. First, the enrollment increase might have occurred through the denominator due to migration if people moved away from SMAs that experience a downturn due to the large increase in imports from China. To examine whether people migrated or not, I used the change in the log of SMA level population size as an outcome. This result is available in Table A7. I failed to find significant migration patterns between 2000 and 2015.

Second, one may worry that the enrollment increase could be due to the rise in the number of colleges and universities when the government loosened the restriction on college establishments in the 1990s (Figure A3). However, this would not lead to the main results. Between 2000 and 2015, while the number of 4-year universities increased from 191 to 201, the number of 2-year colleges decreased from 158 to 138. As shown in Table 2, I found an increase in both institutions. Moreover, the gender difference in enrollments by institution implies that there might be other factors than the large supply of post-secondary education. If a considerable increase in the supply were the reason for our estimates, they would be less likely to be heterogeneous by gender. Finally, I checked whether college education supply correlates to the regional import penetration measure. Figure A4 shows changes in college's admission capacity[28] and import penetration ΔIP_{pt}^{co} across provinces during the sample periods. Although this figure shows province-level correlation, not SMA-level, I failed to find significant correlations between the capacity (supply of college education) and import penetration. Thus, this rules out the possibility that an increase in admission capacity leads to more college enrollments in regions more affected by Chinese imports.

Finally, as a placebo test, I estimated the effect of the import competition on local high school enrollment rates. Since almost all middle school graduates enrolled in high school before Korea increased imports from China,[29] there must be minor effects on high school enrollments. Otherwise, our estimates of college enrollment rates must be due to unobserved factors in the more affected regions. Using the Korean population and housing census, I found that one standard increase in the import penetration in high school enrollments by 35 students per 100,000 residents,[30] which is much smaller than the effect on college enrollments (124.78). This supports that our estimates for college enrollments do not come from unobserved regional characteristics.

[28] To increase or decrease the capacity, colleges need permission from Ministry of Education in Korea.

[29] For instance, in 1991, 97.5% of middle school graduates enrolled in high school (Educational statistics in Korean Educational Development Institute).

[30] The standard error for this is 68.16.

- **Balance check at regional level and randomization inference test**

As a robustness check, I estimated the correlation between the instrument ΔIP^{co}_{it} and pre-determined regional characteristics, using specification (1). To be viewed as quasi-random variation, the instrument should not be correlated with pre-determined variables. Table A8 reports the coefficients of the instrument, using pre-determined regional characteristics as an outcome. None of these are significantly different from 0 except the share of the employed. Table A9 reports the coefficient of the instrument on the outcomes in the initial year of each period. Again, it failed to find any statistical significance, which supports exogeneity of our instrument.[31]

In addition, I used a randomization inference test (permutation test) to test whether the main estimates are statistically significant or just because of a small sample size.[32] The process of this test is the following: As I used industry variation of import growth in other countries as the source of exogenous variation, I randomly assigned the import growth to each industry to construct a placebo instrumental variable. And then, I estimate the effect of this placebo instrumental variable on college enrollments. Figure A5 shows the distribution of t-statistics for the placebo effect from 1,000 random assignments. The red line shows the estimated effect using the actual instrumental variable. I find that the p-value is 0.012, which rejects the null hypothesis that the import competition (measured by import growth in other countries) does not affect college enrollments.[33]

[31] I conducted a regional pre-trends analysis. That is, I regress the pre-trend variables, college enrollment rates in 2000-2005 and 2005-2010, on the instrumental variable, ΔIP^{co}_{it} in 2005-2010 and 2010-2015. I did not find any significant correlation between the pre-trends and instrumental variable.

[32] To do this test, I used `ritest` in Stata (Heß, 2017).

[33] The p-value of the permutation test is the proportion of placebo estimates that are equal to or larger in absolute value than the corresponding estimate from the actual data.

1.5 Conclusion

Many papers have studied the effect of import competition from China on various outcomes in different countries. This paper investigated the effect on educational investments, especially college enrollments among new high school graduates in Korea. Due to many possible channels caused by trade, predicting if more or fewer graduates attend colleges is ambiguous. On the one hand, since the benefit of a high school degree becomes lower due to large imports of low-skilled goods, more graduates enroll in colleges. On the other hand, in an imperfect credit market where students cannot borrow as much money as they want, fewer students can afford a college degree due to a negative income shock.

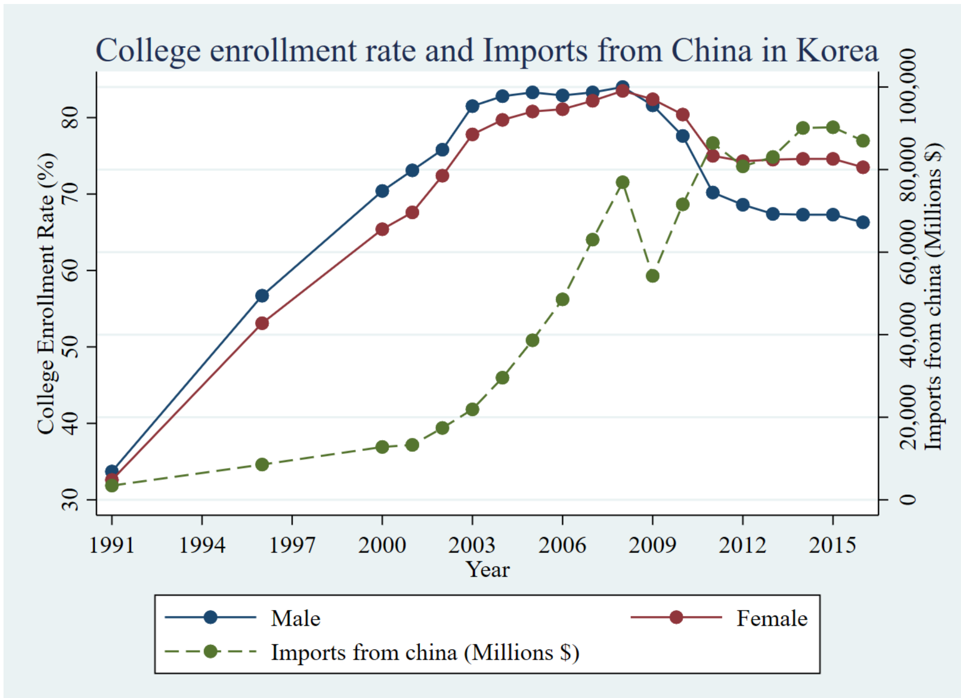
Using administrative statistics of high school graduates from the Office of Education between 2000 and 2015, I found that one standard deviation increase in the local import penetration raised college enrollment by 10% of the average enrollment rate. Moreover, a rise in college enrollment was observed among both male and female students. However, following the import shocks from China, the enrollment patterns were significantly different across genders: while 60% of male students enrolled at 4-year universities, only 23% of female students enrolled at such institutions. Overall, the increase in enrollment rates implies that attending a college involves lower costs and considerable benefits for students. Consistent with this implication, the employment of low-educated workers was more negatively affected by import competition, which suggests that the opportunity cost of attending college decreased, but the college premium increased. In addition, the effect of increased import on low-educated workers was concentrated on men rather than women, which is consistent with the lower educational investments among women I found. Finally, applying the regional sex ratio at birth to measure the negative social norms toward working women, this paper found that women in traditional areas did not enroll at 2-year colleges. Negative social norms toward working women in Korea can be another reason for the moderate increase in educational investment as female students perceived a lower college premium due to the traditional norms.

This paper has two implications. First, while gender gaps in education have narrowed these days in many countries including Korea, import shocks from China do not seem to contribute to these trends. Moreover, as the Korea-China Free Trade Agreement was enacted in 2014, the current trend of rising college enrollment will persist, leading male students to invest more in their human capital than female students. Second, although this paper presented more negative impacts of import competition on male employments, which corresponds to the findings in other countries (e.g., Benguria and Ederington (2021) in Brazil and Besedes et al. (2021) in the US), the long-term effects on labor markets can differ as male students tend to invest in their human capital more than female counterparts to mitigate the negative impacts on their labor outcomes.

An important question that remains unaddressed in this paper is what university students choose to study in response to import competition. For example, they can choose majors related to manufacturing to improve their skills and compete against the workers in exporting countries, or they can select another major to avoid competition. Therefore, by investigating students' choice of the college major after a surge in Chinese import, a more detailed insight can be gained into why students pursue a college education. Gender differences in the choice of the college major (Zafar, 2013) can provide a clearer answer as to why male and female individuals responded differently to import competition with China in terms of their education. I leave this for future research.

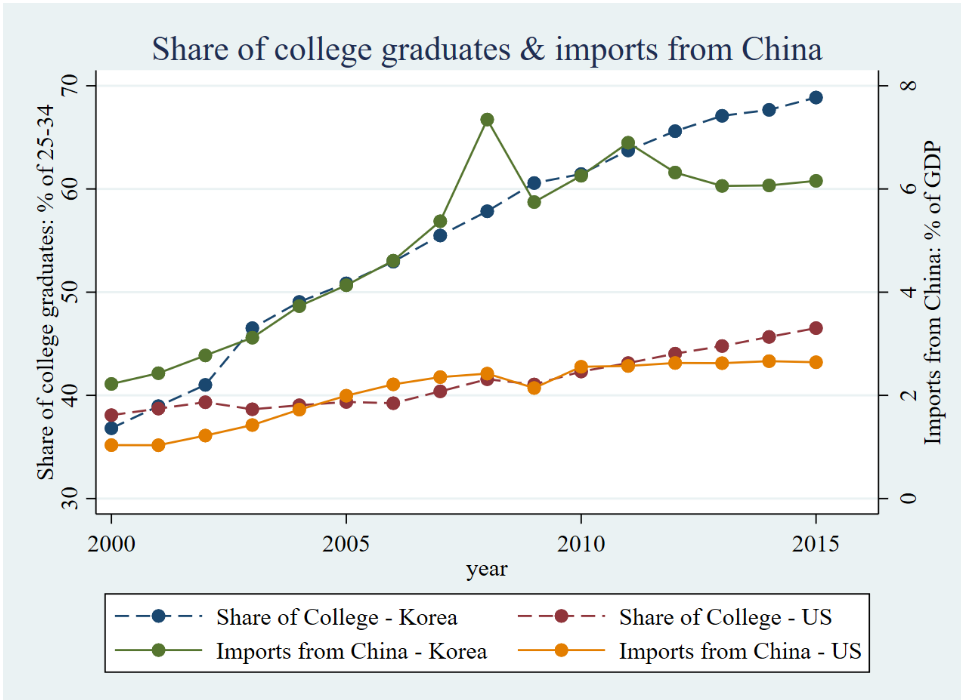
1.6 Figures and Tables

Figure 1: % of college graduates and imports from China (% of GDP) in Korea and the US



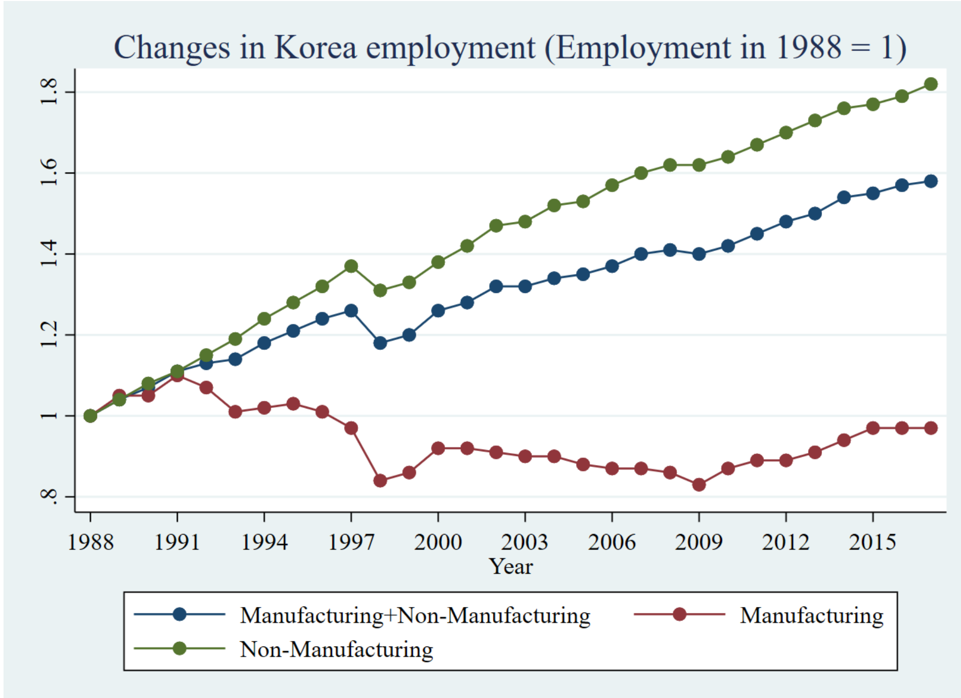
Note: Blue and red lines represent the ratio of college graduates to people aged between 25 and 64. Blue and red dotted lines shows the ratio of import values from China to GDP. Data source: Share of college graduates - OECD. Import values from China - IMF. GDP - World Bank.

Figure 2: College enrollment rate by gender and imports from China in Korea



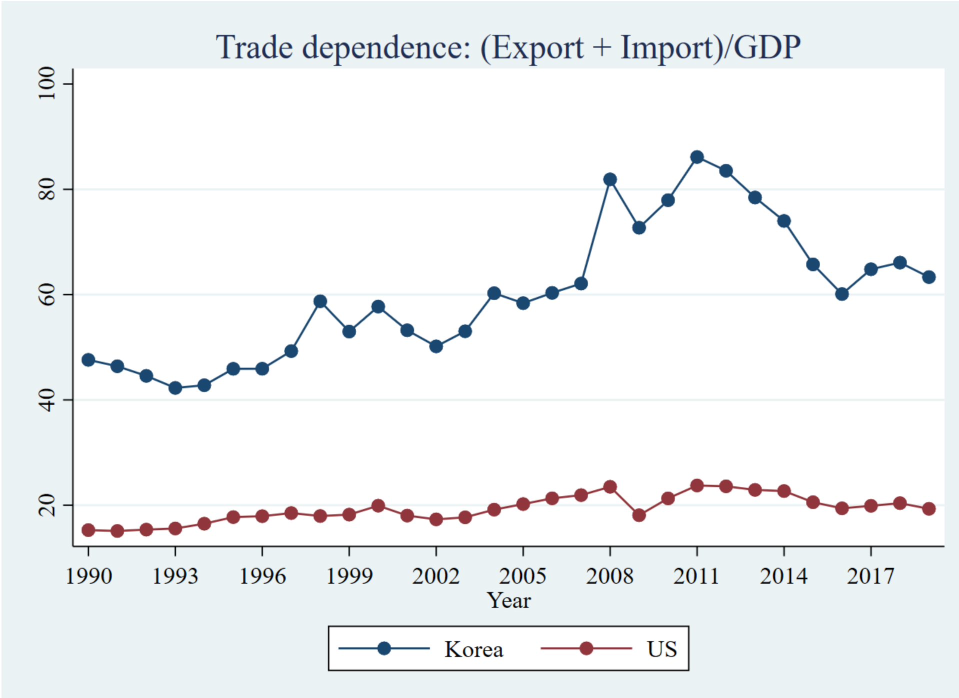
Note: Blue and red dashed lines present the ratio of students who enrolled in a college to high school graduates. A green line shows imports from China in Korea (in millions dollars). Data source: College enrollment rate – Statistics Korea. Import value from China – Korea International Trade Association.

Figure 3: Changes in Korea manufacturing and non-manufacturing employment



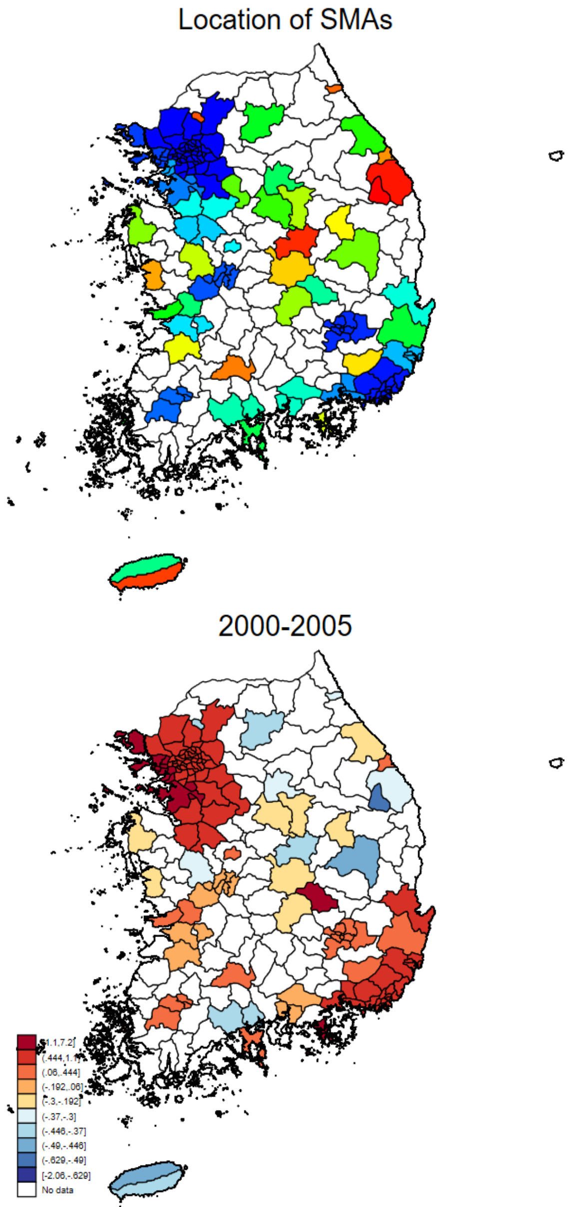
Note: Figure 3 shows the growth of employments compared to 1988 across industries. Employment counts are normalized to 1 in 1988. Data source: Economic activity survey.

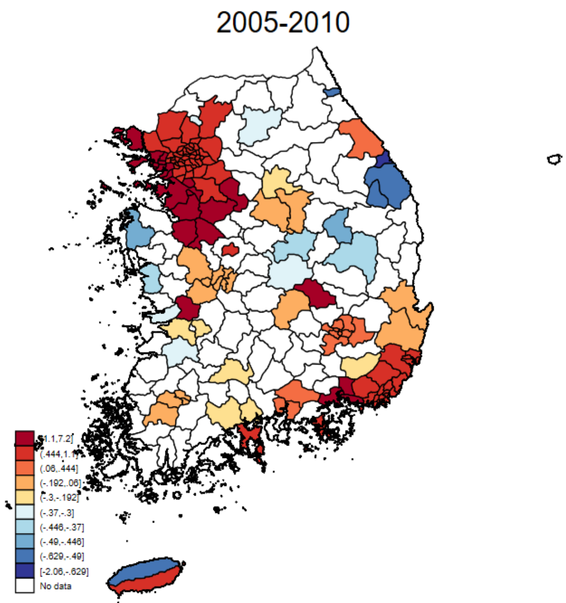
Figure 4: Trade dependency in Korea and the US

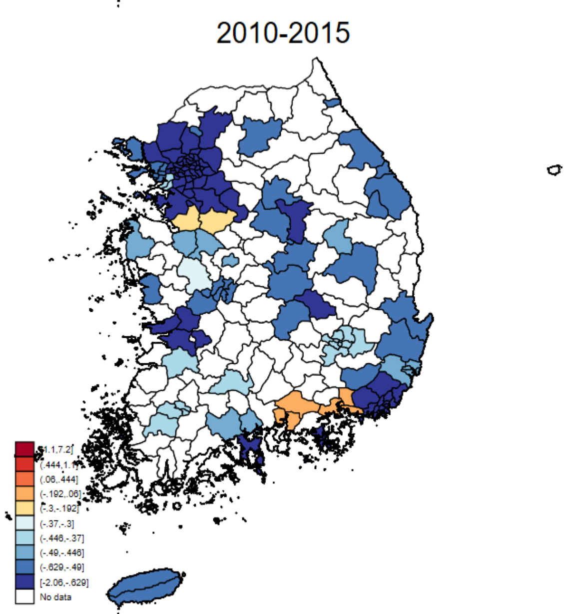


Note: Figure 4 shows the trade dependency in Korea and the US. A measure of trade dependency is $\frac{\text{Imports+Exports}}{\text{GDP}} \times 100$. Data source: OECD.

Figure 5: Location and distribution of local import penetration by SMAs







Note: Local import penetration calculated with the import value of high-income countries.

Data: Census on Establishments (Local labor shares) and UNcomtrade (Import values).

Figure 6-a: Scatter-plot of local import penetration using imports from China in other countries against the local penetration using imports in Korea (First stage regression).

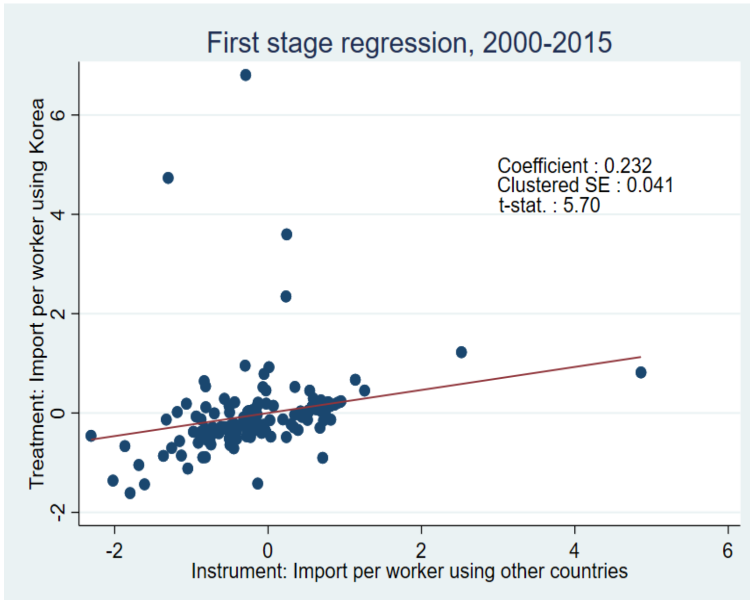
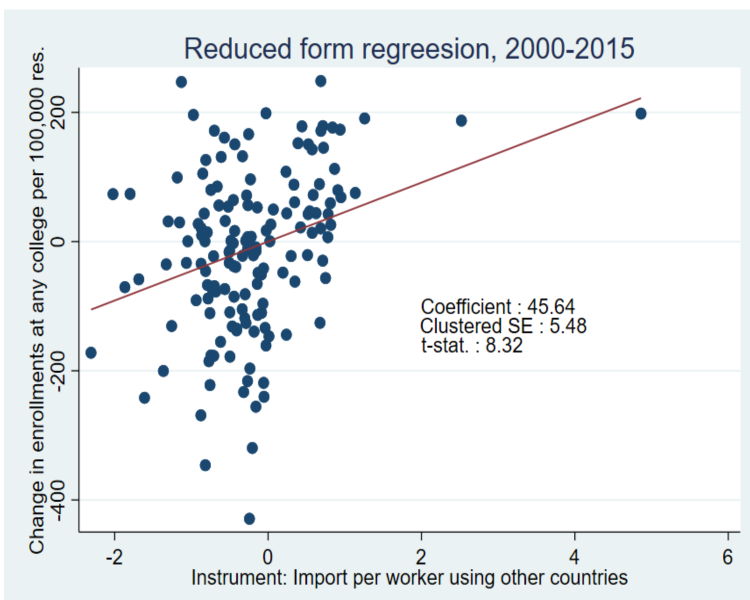


Figure 6-b: Scatter-plot of local import penetration using imports from China in other countries against college enrollments in Korea (Reduced-from regression).



Notes: Figure 6 presents change in import exposure per worker (Figure 6-a) and enrollments at any college per 100,000 residents across an instrument (N = 150). The added variable plots control for lagged manufacturing shares and period fixed effects. Regression models are weighted by the population size in initial year of each period at the regional level.

Table 1-a: Descriptive statistics: outcomes.

Year: 2000, 2005, 2010, 2015				
Per 100,000 resident	Mean	SD	Min	Max
College enrollment (All)	989.74	205.29	605.55	2072.49
2-yr college enrollment (All)	306.57	72.24	109.88	900.67
4-yr univ. enrollment (All)	683.17	175.58	309.12	1446.31
College enrollment (Male)	515.22	118.81	288.14	1098.25
2-yr college enrollment (Male)	156.66	41.12	44.77	462.59
4-yr univ. enrollment (Male)	358.56	98.79	155.55	818.91
College enrollment (Female)	474.53	90.50	213.06	974.24
2-yr college enrollment (Female)	149.91	36.61	65.11	500.50
4-yr univ. enrollment (Female)	324.62	80.20	63.52	627.40
N	200			

Source: Administrative statistics of high school graduates from the Office of Education.

Table 1-b: Descriptive statistics: regional characteristics in the initial year of periods.

Year: 2000, 2005, 2010	Mean	SD	Min	Max
% female	0.51	0.01	0.48	0.54
% 15 to 29	0.21	0.03	0.08	0.26
% Employed	0.43	0.04	0.36	0.60
% college degree	0.23	0.06	0.07	0.32
% manufacturing emp.	0.18	0.09	0.01	0.52
Import penetration using Korea imports	0.97	0.94	-1.15	8.89
Import penetration using Other countries' imports	5.11	6.46	-6.27	44.30
N	150			

Source: Demographic variables are from Korean population and housing census 2 percent sample. Employment share variables for the import penetration are from census on establishments. Industrial imports are from UNcomtrade.

Table 2: Enrollment effect of local import penetration from China.

Panel A	At any-college	At 2-year college	At 4-year university
ΔIP_{it}^{ck}	106.8***	61.18***	45.65
	(40.61)	(23.39)	(28.41)
Mean in 2000	1110.47	348.62	761.85

Panel B	Male	Female	Male	Female	Male	Female
ΔIP_{it}^{ck}	55.14***	51.69**	21.34*	39.85***	33.81**	11.85
	(20.69)	(20.60)	(11.26)	(14.34)	(16.10)	(14.32)
Mean in 2000	596.51	513.96	188.46	160.15	408.04	353.81

N: 150 = 50 (Number of regions) × 3 (Number of periods)

1st stage F-statistic: 11.75

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Numbers in parentheses are standard errors are clustered at the regional level. The coefficients represent the effect of one standard deviation increase in ΔIP_{it}^{ck} on outcomes. All observations are weighted by the population size in initial year of each period at the regional level. Control variables are province-specific period fixed effects, period-specific lagged manufacturing employment shares, the share of females, and the share of people aged between 15 and 29 in the initial year of each period.

Table 3: Employment effect of local import penetration from China across industries.

Panel A	Overall		Manufacturing		Non-manufacturing	
ΔIP_{it}^{ck}	-0.029***		-0.019**		-0.010	
	(0.011)		(0.009)		(0.008)	
Mean in 2000	0.366		0.083		0.282	
Panel B	Male	Female	Male	Female	Male	Female
ΔIP_{it}^{ck}	-0.027***	-0.002	-0.017**	-0.002	-0.009*	-0.001
	(0.009)	(0.004)	(0.008)	(0.002)	(0.005)	(0.004)
Mean in 2000	0.221	0.145	0.058	0.025	0.163	0.120

N: 150 = 50 (Number of regions) × 3 (Number of periods)

1st stage F-statistic: 11.75

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Numbers in parentheses are standard errors are clustered at the regional level. The coefficients represent the effect of one standard deviation increase in ΔIP_{it}^{ck} on outcomes. All observations are weighted by the population size in initial year of each period at the regional level. Control variables are province-specific period fixed effects, period-specific lagged manufacturing employment shares, the share of females, and the share of people aged between 15 and 29 in the initial year of each period.

Table 4: Employment effect of local import penetration across education level.

	All	Male	Female
Overall	-0.010**	-0.012**	0.002
	(0.005)	(0.006)	(0.006)
Mean in 2000	0.642	0.399	0.243
High school or less	-0.019**	-0.016*	-0.002
	(0.008)	(0.009)	(0.007)
Mean in 2000	0.589	0.387	0.202
With any college degree	-0.001	-0.012	0.012
	(0.008)	(0.014)	(0.013)
Mean in 2000	0.721	0.419	0.302

N: 150 = 50 (Number of regions) × 3 (Number of periods)

1st stage F statistic: 11.654

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: All samples are from Korean population and housing census 2 percent sample, have been restricted to males who age 25-34 and females who age 23-32, have been restricted to individuals who live in the same regions 5 years ago. Additionally, High school or less includes all individuals with a high school degree and those below that level. Any college consists of all individuals with 2-year college or 4-year university degrees. Numbers in parentheses are standard errors are clustered at the regional level. The coefficients represent the effect of one standard deviation increase in ΔIP_{it}^{ck} on outcomes. All observations are weighted by the sampling weight of the survey. Control variables are province-specific period fixed effects, period-specific lagged manufacturing employment shares, the share of females, and the share of people aged between 15 and 29 in the initial year of each period.

Table 5: Heterogeneous enrollment effect of local import penetration: Sex ratio at birth

Panel A	At any-col.		At 2-year col.		At 4-year univ.	
	Male	Female	Male	Female	Male	Female
ΔIP^{co}_{it}	13.21***	12.38***	5.11**	9.54***	8.10**	2.84
	(4.34)	(4.36)	(2.53)	(2.77)	(4.01)	(3.83)
Panel B						
ΔIP^{co}_{it}	14.56***	13.83***	6.20**	10.22***	8.36**	3.61
	(4.18)	(4.06)	(2.32)	(2.57)	(4.14)	(4.20)
$\Delta IP^{co}_{it} \times$ Demographic Ind.	-10.92*	-7.77	-7.04*	-12.65***	-3.88	4.89
	(6.42)	(5.88)	(3.80)	(4.66)	(6.47)	(5.02)
Demo. Ind.	11.13	12.00*	8.99	5.52	2.15	6.47
	(7.34)	(6.19)	(5.74)	(4.66)	(4.83)	(5.41)
Mean in 2000	596.51	513.96	188.46	160.15	408.04	353.81

N: 150 = 50 (Number of regions) $\times$ 3 (Number of periods)

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Demographic indicator equals to one if the region reports higher than the median value of the sex ratio at birth in 2000. Numbers in parentheses are standard errors are clustered at the regional level. The coefficients represent the effect of one standard deviation increase in ΔIP^{co}_{it} on outcomes. All observations are weighted by the population size in initial year of each period at the regional level. Control variables are province-specific period fixed effects, period-specific lagged manufacturing employment shares, the share of females, and the share of people aged between 15 and 29 in the initial year of each period.

Chapter 2

Does a Pension Scheme Crowd Out Children? Evidence from Rural China

2.1 Introduction

Children have long been viewed as a source of support for their parents, especially in societies with less developed financial and insurance markets (Caldwell, 1976). Therefore, parents may consider having children as an investment that could yield some degree of security in their old age. However, the family's economic role has consistently decreased as the market, and the state has supplemented or replaced traditional family functions (Lundberg and Pollak, 2007). The provision of support for aged parents is no exception in this regard. Many governments worldwide introduced pension systems so parents have an alternative to their children's support during retirement. This suggests they have less incentive to have and invest in children.

This paper seeks to determine whether and, if so, how parents in rural China reacted to the introduction of a pension program in 2008, the New Rural Pension Scheme (NRPS), regarding the quantity and quality of children. Suppose the financial security of parents is achieved at the cost of children. In that case, a pension program can be a barrier to economic development and growth due to reducing future generations' number and educational level. This question is critical given that many children in rural China could be affected. In rural China, 86.2% of the non-working population aged 60 and above depended on family support before introducing a formal pension system (Wang, 2006). If parents sign up for the pension program and are motivated to invest less in their children, many children can be negatively affected.

Moreover, it is crucial to quantitatively evaluate the pension program since its effect on children is theoretically ambiguous. Considering introducing a pension scheme as lowering the price for future incomes, income and substitution effects affect children in opposite directions. Thus, estimating the impact of the NRPS on children will generate empirical evidence concerning the direction of the effect.

To determine how working-age parents react to the pension program, we use the difference-in-differences approach by exploiting the timing variation of the NPRS roll-out across various counties in China. Using an indicator of the NRPS roll-out as an instrumental variable for the NRPS enrollment, we find that parents who signed up for the NRPS increased their investment in children's education. The increase in educational expenditure was limited to sons, and we failed to find the increase in their spending for daughters' education. When parents increased their expenditures after participating in the pension program, they increased their expenditures related to transportation and extracurricular tutoring, which are easily manipulable. In addition to spending, parents may respond to the new pension program through fertility. Although we failed to find a significant effect on the probability of giving birth to a baby, we found that parents enrolled in the NRPS were more likely to give birth to a boy.

This paper makes three contributions to the literature. First, it helps explain how parents allocate resources to their children concerning the life-cycle model. In terms of inter-generation resource allocation, many prior studies (e.g., Oliveira 2016; Ho 2019) have considered how adult children transfer changes in response to their parent's choices. However, we focus on working-age parents to determine how they alter the resource allocation for their children when the government provides an alternative means of achieving security in old age. The most similar paper is Shen et al. (2020), which analyzed the effect of NRPS on fertility. Compared to Shen et al. (2020), this paper corrects biases in estimates of the impact of NRPS by using NRPS roll-out as an instrumental variable and data for longer periods. Thus, while they find a decrease in fertility, this paper fails to find a significant effect on fertility. As we found the negative impact on fertility when we did not correct the biases, we suspect their results would be biased due to the endogeneity of the NRPS enrollment.

Second, this paper contributes to the policy evaluation of the NRPS.[1] Although most prior studies have evaluated the impact of the NRPS on the elderly, this paper evaluates its impact on children, which is similar to the approach of Huang and Zhang (2021) and Yang and Chen (2022). These papers show that children in the NRPS counties receive more pocket money and care from grandparents, are more likely to stay in school, and children living with grandparent pensioners are healthier. However, this paper differs from such works in two ways: First, it specifies the more critical impact channel on children, namely parents rather than grandparents. Considering that children tend to live with their parents and are more affected by their parents than their grandparents, it is more important to understand how parents react to the new pension policy. Second, this paper investigates outcomes that the earlier two papers did not consider, namely fertility and educational expenditure. Thus, the present paper complements previous studies by widening our understanding of the impact of pension programs on children.

Third, this paper contributes to the literature on how policies affect cultures. Specif-

[1] See Yang and Chen (2022) for a summary of the literature concerning the effects of the NRPS on various outcomes.

ically, this paper explores how a pension program changes son preference. China has a strong son preference, which is even stronger in rural areas. Researchers have pointed out that sons traditionally provide old-age support for their parents in China, which would be one reason for the strong preference (Jayachandran, 2015). If the government provides a pension program that can play the role of sons, parents would consider investing more in their daughters, reducing gender gaps. Unlike this prediction, we find that parents who enrolled in NRPS were more likely to have a son and invested more in their son.

The remainder of the paper is organized as follows. Section 2 explains about the NRPS with details. Section 3 provides theoretical backgrounds to analyze parents' utility maximization problem when introducing a new pension scheme. Section 4 specifies our empirical strategy. In section 5, we describe the data and variables used in this analysis. Section 6 presents the main results and evidence, including validating the assumptions of our empirical strategy. Finally, section 7 concludes this paper.

2.2 New Rural Pension Scheme

As the population of rural areas ages due to the one-child policy, migration of young adults to urban areas, and increasing life expectancy, the welfare of the elderly in rural areas is becoming a critical issue for policymakers in China. In the past, there have been few policies to assist the elderly, such as the voluntary contributions to the Rural Old-age Pension Program initiated by the Ministry of Civil Affairs in 1991 and terminated in 1999 (Zhang et al., 2019). Thus, it is more common for the rural elderly to live poorly, without any social safety net, depending on financial support from their children (Wang, 2006).

New Rural Pension Scheme (NRPS) was initiated in 2008 and announced in 2009[2]

[2]Stepan and Lu (2016) explained the process of introduction of the NRPS in detail. Based on the pilot project initiated in such as Baoji City in 2008, on 1 September 2009, the State Council published a guideline for the new rural social insurance pension. According to Reuters, 56 million rural residents

to deal with the problem of poverty, especially among the rural elderly. When rolling out the NRPS to rural counties, the government documents indicate that the central government randomly chose some counties for early implementation and introduced the NRPS gradually (State Council of China, 2009).[3]

Figure 1 explains how the NRPS is financed and provides pension benefits (Zhao et al., 2016). All rural residents aged 16 or older who are not in school and not enrolled in an urban pension scheme are eligible to enroll in this program. Although the NRPS is similar to the previous program based on voluntary participation, the main difference in the NRPS is that the government provides voluntary enrollees with subsidies and basic pensions. Once a decision is made to participate in this program, an individual chooses from five categories of premium: 100 (around 17 USD), 200, 300, 400, and 500 RMB per year per person. The central and local governments provide enrollees with subsidies: the minimum matched subsidy amount should be no less than 30 RMB per person per year for the first 100 RMB of individual premiums with a maximum amount capped at 50 RMB (Zhang et al., 2019). The premium with the matched subsidy is accumulated at a 1-year deposit interest rate in the individual account.

Participants who contribute for at least 15 years will be eligible for benefits from the program at age 60. There are two sources of the NRPS benefits: a) non-contributory (basic) pension and b) individual contributory pension account. The basic pension is fully financed by the government and is available to all participants when they become 60. Basic pension levels vary across provinces, with a minimum payment of 55 RMB per month and a maximum payment of up to 360 RMB (around 60 USD) for some wealthier

had joined this program by the end of 2008, whereas 5.12 million rural residents received total pension payments of 5.68 billion RMB in 2008.

[3] While Cheng et al. (2018) exploit regional variation of the announcement timing, Xi et al. (2019) use the variation of the implementation timing in each county to estimate the effect of the NRPS on outcomes. We follow Xi et al. (2019) and construct a binary variable for the program roll-out in a county and related variables. See section 5.3 for more about this.

provinces (e.g., Shanghai, Beijing). The basic pension is provided until the participant dies. The accumulated contribution in the individual account will be paid back to the pensioner for 139 months starting at 60.[4]

For instance, if an individual has chosen the lowest premium (100 RMB≈17 USD/year), the pension received from the contributory and non-contributory accounts until turning 71.5 years old will be around 20 and 55 RMB per month, respectively. The individual will receive about 75 RMB per month, about 40% of the national poverty line (Zhao et al., 2016). After turning 71.5 years old, the individual will get only 55 RMB per month as the basic pension. We consider the introduction of the NRPS as a reduction in the price of retirement benefits since individuals receive subsidies when they pay for premium and non-contributory pensions from the government after they become 60 years old.

Although the government provides subsidies for premiums and basic pensions starting at 60 years old, many people did not know about the NRPS. In the 2011/12 Chinese Longitudinal Healthy Longevity Survey, more than 50% of the non-pensioners did not know about the NRPS (Cheng et al., 2018).

2.3 Theoretical Framework

In this section, we borrow Oliveira (2016)'s theoretical framework with a slight modification to analyze the effect of the NRPS on fertility (the quantity of children) and human capital investment in children (the quality of children). The model assumes two periods, childbearing years and retirement years, to employ a life cycle perspective. The model also assumes for simplicity that children do not affect utility in the second period. This may be true, for instance, since adult children live separately from their parents and make

[4]139 months are related to China's life expectancy in 2009, which was 71.5 years old. When they become 60 years old, they start receiving a pension from the individual account, and this pension ends when they get 71.5 years old, which is 139 months after 60 years old.

decisions independently. Finally, we assume that the marginal utility of all components is infinity at 0 and always positive so that parents consume at least a positive amount, and consuming more assures higher utility.

If parents raise a child to get support from them during retirement, and parents do not get any utility directly from children, then a pension is a perfect substitute for children. In this case, the model becomes

$$\max_{c_1,c_2,Q,s} u(c_1) + u(c_2) \tag{2.1}$$

$$\text{subject to } c_1 + \pi \cdot Q + p \cdot s = I$$

$$s + Q = c_2,$$

where π and p are the marginal cost per-child and price of the pension, respectively. Parents choose Q (the quantity and quality of children), and the amount of pension s to be consumed in the second period. In this scenario (Old-age security model), parents either choose to have (and invest in) children ($Q > 0, s = 0$) or a pension ($Q = 0, s > 0$) based on which is the cheaper option. For instance, if π is smaller than p due to the high earning potential of children, then the parents choose to invest in children without investing in a pension, and vice versa.

The introduction of the NRPS can be interpreted as lower $p^{new} < p^{old}$ since this program provides a basic pension and subsidy to enrollees in addition to the accumulated premium they have paid. When the NRPS lowers the price from p^{old} to p^{new}, some households whose $\pi \in [p^{new}, p^{old}]$ enroll in the NRPS to substitute pensions for children. Thus, the effect of the implementation of the NRPS on fertility and expenditures for children is expected to be negative. In contrast, a household with π is in $[0, p^{new})$ or $(p^{old}, \bar{p}]$,[5] there is no effect of the implementation on Q because the new price is still expensive compared to π if $\pi \in [0, p^{new})$, or parents already spent all on saving before

[5] $\bar{p}$ is an upper bound of p.

the introduction if $\pi \in (p^{old}, \bar{p}]$.

On the other hand, if parents raise children and spend money on them purely for their utility, considering children to be a consumption good in the first period, the model becomes

$$\max_{c_1, c_2, Q, s} u(c_1, Q) + u(c_2) \tag{2.2}$$
$$\text{subject to } c_1 + \pi \cdot Q + p \cdot s = I$$
$$s = c_2.$$

In this scenario (Altruism model), the effect of the implementation is ambiguous. On the one hand, as the price of consumable goods in the first period (c_1, Q) becomes more expensive than s, parents spend less on their children. On the other hand, the pension becomes cheaper, so they consume more for their children through the income effect.

Finally, if parents raise children to get utility and support directly from them in the second period (Altruism&Old-age security model), parents solve the following problem:

$$\max_{c_1, c_2, Q, s} u(c_1, Q) + u(c_2) \tag{2.3}$$
$$\text{subject to } c_1 + \pi \cdot Q + p \cdot s = I$$
$$s + Q = c_2.$$

Parents whose $\pi < p$ do not invest in saving $(Q > 0, s = 0)$ because investing in children is cheaper for c_2 in addition to getting utility directly from Q. Thus, those whose $\pi \in [0, p^{new})$ will not change their consumption as the new price for saving is still higher than π. This is the same as in the Old-age security model. If $\pi \geq p$, parents invest in both saving and their children since saving becomes a cheaper way to prepare for c_2. For parents whose $\pi \in [p^{new}, \bar{p}]$, there must be a substitution and income effects like the

Altruism model, and the effect is ambiguous.[6]

To sum up the three models, it was determined that only the Old-age security model could predict negative effects on children if parents enroll in the NRPS, while the effect is ambiguous in the other two models. Table 1 summarizes the effects of NRPS implementation on Q based on the models and the location of π relative to p^{new} and p^{old}.

2.4 Empirical Strategy

To investigate the effect of enrollment in a pension program, we estimate the following model using panel data:

$$Y_{ict} = \beta_1 D_{ict} + \beta_2 X_{ict} + v_c + \gamma_t + \epsilon_{ict} \tag{2.4}$$

where i, c, and t are the indexes for household, county, and year, respectively; Y_{ict} denotes outcome variables, such as fertility and educational expenditures for children; D_{ict} is a binary variable that is equal to 1 if a household i enrolls in the NRPS in year t and 0 otherwise; and X_{ict} is a vector of household time-variant characteristics.[7]

β_1 indicates the effect of NRPS enrollment on the outcome, a parameter of main interest. Despite controlling for household and county characteristics through fixed effects and X_{ict}, D_{ict} might be endogenous because a household can voluntarily join the NRPS. For instance, a household with fewer children is more likely to participate in the NRPS as they have fewer children who will support their parents. Ebenstein and Leung (2010) showed that parents without a son are more likely to enroll in the Rural Old-age Pension Program, the pilot program that was terminated in 1999. This implies that β_1 may be

[6] Altruism&Old-age security model is similar to the model in Shen et al. (2020). However, since Shen et al. (2020)'s model assumes a quasi-linear utility function, the predicted effect of NRPS on the investment for children is negative, which would be false if there is an income effect.

[7] See section 5.4 for more details on each variable.

biased due to simultaneity. Finally, if parents consider educational expenditure as saving, unobserved preferences for saving could cause omitted variable bias (Poterba et al., 1995). Table A1 compares household characteristics with children depending on whether they enroll in the NRPS or not. Parents with the NRPS are older than those without the pension, and fathers with the NRPS are more educated than those without the NRPS.

To deal with the endogeneity of D_{ict}, we exploit the regional variation of the implementation of this program over counties. Specifically, we use an indicator of NRPS implementation as an IV, Z_{ct}. This setting is similar to Imbens and Angrist (1994), which estimated the local average treatment effect (LATE) by a random assignment with imperfect compliance. The LATE can be estimated by a ratio of the estimated intent-to-treat (ITT) effect and the proportion of compliers.

Given county (household) and year fixed effects on the right-hand-side of (4), this empirical strategy is difference-in-differences (DID) with IV, which is similar to Duflo (2001). In this context, the ITT and the proportion of compliers can be estimated by comparing the corresponding changes in outcomes and the proportion of enrollees in the NRPS over time between program and non-program counties. The ratio of these two estimates is the effect of NRPS enrollment on the outcomes, β_1, according to the LATE theorem.

Specifically, the equations for the first-stage and reduced forms are shown below.

$$\begin{aligned} D_{ict} &= \gamma_1 Z_{ct} + \gamma_2 X_{ict} + \gamma_c + \gamma_t + \epsilon_{ict} && \text{First stage regression} \\ Y_{ict} &= \alpha_1 Z_{ct} + \alpha_2 X_{ict} + \alpha_c + \alpha_t + \tilde{\epsilon}_{ict} && \text{Reduced form regression.} \end{aligned} \tag{2.5}$$

These equations are related to the theoretical analysis in the previous section. α_1 in (5) provides the effect of NRPS implementation on children ($= \frac{\partial Y}{\partial Z}$). The sign of α_1 is ambiguous due to the opposite directions of the income and substitution effects analyzed in Table 1. On the other hand, the impact of the implementation on the enrollment

($= \frac{\partial D}{\partial Z}$) is trivially expected to be positive, which is γ_1 in (5). If we get $\frac{\partial Y}{\partial Z}$ and $\frac{\partial D}{\partial Z}$, we can infer the effect of enrollment in the NRPS on outcomes by calculating ratio ($\frac{\frac{\partial Y}{\partial Z}}{\frac{\partial D}{\partial Z}} = \frac{\partial Y}{\partial D}$). This is equivalent to estimating the effect of D on Y using Z as an instrumental variable (IV) for D. The underlying assumptions for the ratio to be interpreted as a causal effect are the non-zero correlation, common trend, and exclusion restriction assumptions.

In this case, these three assumptions are more likely to hold. As people in rural areas could not sign up for the NRPS before its roll-out, the coefficient of the program implementation's indicator in the first stage regression must be positive (non-zero correlation). This assumption is testable, and this is supported by Figure 2 and the result of the first stage regression in section 6.1. Second, in the absence of the program, the change in fertility and expenditure for children's education would not have been systematically different in the program and non-program counties under the common trend assumption. To test this assumption, we estimate pre-trends on fertility and the effect of the NRPS on expenditure with urban samples in section 6.4. We do not find any evidence against the common trend assumption. Finally, in this context, the exclusion restriction assumption means that the roll-out of the NRPS should only affect fertility and investment in children through parents' decisions on pension enrollment. This condition is not testable, but this would also hold because although the roll-out of the NRPS in a county cannot change anything for parents, it allows them to enroll in the pension scheme, which means that the roll-out is likely to affect the outcomes through NRPS enrollment. Thus, we interpret an estimate of β_1 using the IV method as a causal effect of NRPS enrollment.

Finally, we compute the standard errors clustered at the county level. This is appropriate because the level of treatment is at the county level (Abadie et al. 2022).

2.5 Data

2.5.1 Data Introduction

China Family Panel Studies (CFPS) is a nationally representative, annual longitudinal survey of Chinese communities, families, and individuals launched in 2010. The studies focus on the economic, as well as the non-economic, wellbeing of the Chinese population, with a wealth of information covering such topics as economic activities, education outcomes, family dynamics and relationships, migration, and health. In the 2010 baseline survey, the CFPS successfully interviewed almost 15,000 families and almost 30,000 individuals within these families, for an approximately response rate of 79%. The 2012 follow-up wave successfully resurveyed more than 85 percent of the 2010 baseline sample.[8] Basically, we use the 2010 and 2012 waves of this data.

2.5.2 Outcome Variables

This study aims to consider parents' reactions to a pension program on the quantity and quality of children. To estimate the effect of the NRPS on quantity, we use an indicator of giving birth to a baby in year t in a household i.

- $Birth_{i,c,t}$ is equal to 1 if household i gives birth to a baby in year t and 0 otherwise.

As CFPS has birth data of each child under 15 years old, it is possible to construct $Birth_{i,c,t}$ from 1995 to 2010 with the 2010 wave and from 1997 to 2012 with the 2012 wave. For instance, if a family reports that they have a 5-year-old boy and a 3-year-old girl in the 2012 wave, $Birth_{i,c,2007}$ and $Birth_{i,c,2009}$ are equal to 1, and $Birth_{i,c,t}$ is 0 for the other t. We use $Birth_{i,c,t}$ from 2004 to 2012, four years before and after the first roll-out of the NRPS.

[8]https://www.isss.pku.edu.cn/cfps/en/about/introduction/index.htm

To estimate parents' reactions on the quality of children, we use the average educational expenditure per child in a household as a dependent variable.

- $Education_{i,c,t}$: Average of educational expenditure per child in household i at county c in year t.

This represents an educational investment in their children, which would affect their abilities to support retired parents. Unlike the birth outcomes, CFPS does not have yearly data on expenditure but only has available data for 2009/2010 and 2011/2012.[9] Therefore, the estimates on educational expenditure represent an instantaneous effect. Another concern with using expenditure data is the effect of outliers on estimates, as the expenditure distribution is very right-skewed. Thus, we used data censored from the 99th percentile for the primary analysis.[10] Finally, I adjusted all RMB amounts to the 2015 CPI level to address inflation.

2.5.3 The Enrollment of the NRPS and Length of Exposure on NRPS Implementation

A key independent variable is a dummy variable for whether a household enrolled in the NRPS. This data has year and month information as to when a husband and a wife first paid for this program. Based on such information, we constructed the dummy variable from 2004 to 2012.

- $D_{i,c,t}$: a dummy variable that is equal to 1 if either father or mother enrolls in the NRPS in household i at county c in year t and 0 otherwise.

[9]In these waves, we exclude the data surveyed in 2010 and 2012 since the number of observations in these two years is negligible.

[10]Trimming some portion of data is common to exclude outliers. When analyzing the effect of hospital admission on medical spending, Dobkin et al. (2018a) exclude the top 0.05 percent samples in medical expenditures. Lusardi and Mitchell (2007) trim the bottom and top 1% of the wealth distribution to estimate the effect of financial literacy on wealth.

As discussed in the previous section, we use an indicator of NRPS implementation as an IV for $D_{i,c,t}$. A main problem in constructing the IV is that it needs to have information on the initiation date of the NRPS in each county. Following Xi et al. (2019), we thus calculate the earliest date in the data that people paid a premium for and received a basic pension from the NRPS in a county and use the earliest date in a county as the date of the implementation in each county. Then, we define program county as one that had been exposed to the NRPS for more than six months in year t, following Cheng et al. (2018) for two reasons. First, due to the difficulty in understanding the NRPS and distrust from the failure of the last pilot program, it took time to persuade people to enroll in this program. Second, social learning is an important channel to increase program take-up in rural China, which also takes time (Liu et al., 2014). Six months may not be a reasonable threshold to define the program counties. Thus, we use alternative thresholds, namely, 0 months and 12 months, to estimate the effect of NRPS enrollment in section 6.4. The results from using the alternative thresholds are similar to the main results.

Based on this, we calculate the program duration and construct a dummy variable for the program counties, which is as follows:

- $Z_{c,t}$: a dummy variable that is equal to 1 if the $Duration_{c,t} \geq 6$ and 0 otherwise.
- $Duration_{c,t}$ is equal to $(t - y\,NRPS_c) \times 12 + (12 - m\,NPRS_c)$ where $y\,NRPS_c$ and $m\,NRPS_c$ represent the year and month of NRPS roll-out in county c, respectively.[11]

Note that $Duration_{c,t}$ has a measurement error since the memory of people could be wrong, or the earliest date could be different from the initiation date of the NRPS. To

[11] When the announcement date rather than the implementation date is used, there is an measurement error issue as well since this would be different from the exact dates of program initiation for each county (Cheng et al. 2018).

mitigate this measurement error, we only exploit the variation of whether $Duration_{c,t}$ is more or less than six months by using $Z_{c,t}$. However, this variation may not be enough to estimate the effect of this program. Thus, as alternative specifications, we exploit the variation of $Duration_{c,t}$ further by using several dummy variables indicating the duration in years[12] or using $Duration_{c,t}$ in years as a continuous instrumental variable. The results, available in section 6.4, are similar to the main results.

Using the earliest date in a county as the adoption year, Table 2 provides the regional variation of NRPS roll-out timing in the data.

2.5.4 Other Independent Variables

We include many variables on the right-hand side of the regression equation to control for various unobserved characteristics in (4): First, to control for time-variant variables, we include several household characteristics such as the father's and mother's age, their educational level. In addition, we include other characteristics used in previous papers studying household expenditures. Those include the number of children, the number of generations, the age of a child, and the total amount of assets (adjusted value to 2015 level) in a household. All time-variant variables come from the 2010 and 2012 waves of CFPS. However, we only use parents' age as a control variable when estimating the effect on fertility outcomes because all other variables before 2009 are unavailable.

2.5.5 Sample Restriction and Descriptive Statistics

This section summarizes sample restrictions for fertility and educational expenditure analyses. Since the empirical strategy is difference in differences, it needs untreated (before the roll-out of the NRPS) and treated data (after the roll-out of the NRPS) in program counties (treatment group). In comparison, only untreated data are needed for

[12] In this case, the first stage and reduced from regressions are the same as event study design.

non-program counties (control group).

Table A2 describes the sample restriction due to the empirical strategy. As fertility data are available from 2004 to 2012, it is possible to use data from all households. However, only the data in 2009 and 2011 are available for educational expenditure for children. Thus, the data of households in counties that implemented the NRPS before June 2009 must be excluded because they were already treated in 2009 according to the definition of the program county.[13]

In addition to the sample restriction due to the empirical strategy, we exclude samples with urban hukou[14] (non-agricultural hukou) for both analyses, because this policy is for rural areas. Moreover, to remove the effect of receiving a basic pension from NRPS, we exclude households where the age of mother or father is higher than 60 because they are eligible to receive a pension. Finally, for the fertility analysis, we restrict households wherein the mother's age is higher than 16 but lower than 50, which is the range of reproductive age followed by the World Health Organization (WHO). Table 3 summarizes the sample restrictions for each analysis.

Based on the restrictions, we use 5,040 households for the fertility analysis. The average ages of the mothers and fathers in the survey were 32.32 and 30.48, respectively. On average, 9.1% of households gave birth to a baby, and among them 4.8% gave birth to a boy in each year. Table 4 provides summary statistics of the samples for the expenditure analysis. They consist of 2,029 households from 73 counties. Among them, there are 444 households (22%) in 17 counties (23%) categorized as non-program counties according to

[13]If they were included, they must have been used as a control group since there is no variation on $Z_{c.t}$ among them between 2009 and 2011. In this case, as Goodman-Bacon (2021) pointed out, the estimates in (5) would have included difference in differences using the early treated group as a control group, which cannot provide an estimate of the average treatment effect but an estimate of the sum of treatment effects with negative weights for some treatment effects.

[14]Hukou is a system of household registration in China. This is connected to social programs provided by the government, which assigns benefits based on agricultural and non-agricultural residency status.

the definition of the program counties. Looking at these samples by wave, parents spent more on their children's education, and more parents signed up for the NRPS in the latter wave. This could be attributed to the time effect or the income and substitution effects of the NRPS.

2.6 Results

In this section, we report the effects of NRPS using two-stage least squares (2SLS) regressions and the effects of NRPS implementation using reduced-form regressions. While we provide qualitative interpretations (the direction of NRPS enrollment effects) from 2SLS, we primarily offer quantitative interpretations from reduced form regressions due to large standard errors in 2SLS compared to the mean of outcomes.

2.6.1 First-stage Regression

Table A3 and A4 report estimates of the first stage regression in (4) for each outcome. As expected in section 4, the coefficient on program counties is significantly positive. Moreover, we need to check whether the first-stage correlation is strong. Andrews et al. (2019) suggest to use effective F-statistic of Olea and Pflueger (2013) to detect weak instrument. In the estimates, all of the effective F-statistics for the fertility expenditure in section 6.2 are larger than 23, which implies that the maximum bias of our IV estimates must be less than 10% of the bias of OLS. Thus, this supports that the first identification assumption (non-zero correlation) holds here. Figure 3 plots the dynamic effects of NRPS roll-out.[15] As expected, the enrollment rate increases as the counties implemented NRPS earlier. This also gives us confidence in the assumption.

[15]Specifically, the model used here is $D_{ict} = \sum_{l=-4,\neq -1}^{4} \alpha_l \cdot 1\{Dur\_year_{c,t} = l\} + \alpha_c + \alpha_t + \mathbf{X}'_{ict}\alpha_X + \epsilon_{ict}$ where Dur_yearc, t is $Duration_{c,t}$ in years.

2.6.2 Results: Fertility and Educational Expenditure

Tables 5 and 6 list estimates on fertility and the average educational expenditure in a household. In Panel A of each table, column 1 displays an estimate by ordinary least squares (OLS) regression while columns 2∼4 display estimates by two-stage least squares (2SLS) regression with different control variables to correct bias from endogeneity. The coefficients of the instrument on outcome are shown in Panel B, which are the coefficients of the reduced form regression in 2SLS.

Table 5 displays estimates on the indicator of giving birth to a baby each year after enrolling in the NRPS. If the substitution effect is more significant than the income effect of the NRPS, fewer parents give birth to a baby after NRPS enrollment. The coefficient of OLS (Column 1) supports this finding, indicating that those parents with NRPS were 1.4% less likely to give birth compared to those without NRPS. However, as NRPS enrollment might be endogenous, the OLS coefficient could be biased. Accordingly, the coefficient becomes positive after correcting bias from endogeneity by IV. A household with the NRPS was more likely to give birth each year than those without the NRPS. However, they are all statistically insignificant due to the large standard errors. Thus, our IV estimates do not provide meaningful information about fertility. Panel B in Table 5 presents the coefficients of the reduced form regression, which can be interpreted as the effect of NRPS implementation (ITT). Parents in program counties are 0.6 percentage points more likely to give birth compared to those in non-program counties, which is approximately 7% relative to the mean. Similar to the IV estimates, the standard error is too large to be conclusive. Finally, Figure 4 plots the dynamic effects of the implementation on fertility, which are comparable to the estimates in Panel B of Table 5.[16] Although the effects get larger as they are exposed to the NRPS longer, they are statistically insignificant.

While a household needs time to react to the NRPS in terms of fertility as described

[16]The model used here is $Y_{ict} = \sum_{l=-4,\neq -1}^{4} \alpha_l \cdot 1\{Dur\_year_{c,t} = l\} + \alpha_c + \alpha_t + \mathbf{X}'_{ict}\alpha_X + \epsilon_{ict}$.

in Figure 4, they need less time to respond through changes in expenditure. Further, they need to pay a premium immediately after enrolling in the NRPS, which automatically requires changes in expenditure. Therefore, these changes in expenditure are more visible in representing parents' reactions to the NRPS on children in a short time. Table 6 shows the estimates of the effect on average education expenditure for each child. Column 1 indicates that after enrollment in the NRPS, they spent 227 RMB (around 38 USD) more in educational expenditures for children per year than those not enrolled. Moreover, after correcting endogeneity bias, the estimates become much larger. This suggests that the income effect from the NPRS is much larger than the substitution effect. Further, the NRPS caused a significant increase in investment in children's education. Panel B demonstrates that parents in counties exposed to NPRS over six months spent an average of 570 RMB (around 95 USD) more per year on their child's education, which is 47% of the mean. According to the estimates on educational expenditure, the NRPS is a very effective way of improving education levels in rural areas by motivating parents to invest more in their children's education.

2.6.3 Heterogeneous Effects

Which categories of educational expenditure are increased?

It has been stated that parents enrolled in the NRPS spend more on children's education than those not enrolled in the NRPS. In this subsection, we answer the following question: if parents increase educational spending after enrolling in a pension program, which categories do they increase?

To answer this question, (5) is estimated using the following more detailed information about expenditure: tuition, the total cost of books, lodging, extracurricular tutoring, and transportation. Table 7 shows the effect of the NRPS on each subcategory. Parents increased extracurricular tutoring and transportation costs among five categories.[17] It

[17] Although the impact on extracurricular tutoring is not statistically significant in this specification,

should be noted that relative to the former three subcategories, the costs for extracurricular tutoring and educational transportation are more easily manageable for parents. This implies that parents enrolled in the NRPS increase educational spending, which they can change quickly.

Son preference

In section 3, we analyze parental decisions on the quantity and quality of children without considering their gender. However, if parents prefer sons over daughters, they would choose to have more sons and invest more in their education than their daughters. Moreover, China has a strong preference for sons. Researchers have indicated that sons traditionally support aging parents in China, which would be one reason for this strong preference (Jayachandran, 2015). In this situation, Chinese parents in rural areas would react to the NRPS differently based on the gender of the children. In this section, we investigate whether parents responded differently according to the gender of their children by considering the number of male babies and the educational expenditure on sons and daughters.

Table 8 presents the effect of the NRPS on the probability of giving birth to a boy and a girl. According to this table, parents enrolled in the NRPS were more likely to give birth to a boy, while the effect on giving birth to a girl is insignificant. Parents in the program counties were 0.8 percentage points more likely to give birth to a boy than those in the other counties. This is 17% of the average probability of giving birth to a boy. Figure 5 compares the dynamic effects of NRPS implementation on the likelihood of giving birth to a boy and a girl.[18] As parents are exposed to the NRPS longer, they are

they become significant with lower standard errors when using different thresholds to define a program county dummy. However, estimates for tuition, the total cost of books, and lodging are insignificant in any specification.

[18] These estimates are comparable to the estimates in Panel B of Table 8.

more likely to give birth to a boy, while there are no significant effects on the probability of giving birth to a girl.

Table 9 presents estimates of the effect of the NRPS according to gender. Since the gender of a child would be endogenous, we estimate the impact of the NRPS on the first child in a household separately by gender, assuming the gender of the first child is random.[19][20] While parents enrolled in the NRPS spent more on their sons per year than without the NRPS, they also spent more on their daughters per year. However, since the estimate on a daughter is insignificant after controlling for family-fixed effects, it is impossible to conclude the spending for daughters. Parents in the program counties spent 669∼886 RMB (about 112∼148 USD) more on their son, which is about 49∼69% of the average annual spending for a son.

Thus, we fail to find evidence that the roll-out of a new pension program reduces son preference in a household. These results differ from the expectation that the roll-out of NRPS would reduce son preference since adult sons traditionally support old parents in China. Our results are consistent with small substitution effects between NRPS and investment for a son. As parents do not consider NRPS a substitute for their son and NRPS is a cheaper way to prepare for their retirement, significant income effects from NRPS lead them to spend more on their son. The failure to find significant investments for a daughter suggests that there may be other reasons for son preference in China, such as the unique roles of sons in ancestor worship (Jayachandran, 2015). The other possibility is cultural persistence (Giuliano 2021). Bau (2021) found that matrilocal daughters in Indonesia received less education after the government introduced the pension program. She observed a reduction in completing secondary school at least ten years after the roll-

[19] Another paper using this identification assumption is Dahl and Moretti (2008).

[20] In this data, the sex ratio of the first child in all households is 0.509. In contrast, the percentage of a second child and higher parity than the third are 0.553 and 0.616, respectively. Thus, relative to higher parity, the sex ratio of the first child is random.

out. Thus, our finding is consistent with her paper and previous papers showing cultural persistence in the medium term (Giuliano 2021).

2.6.4 Robustness check

In this section, we provide several estimates for robustness check.

Alternative specifications

As alternative specifications, we exploit regional variation of the implementation timing further: First, we divide the program county dummy into several dummy variables indicating years relative to NRPS implementation[21] and use these as IV. Thus, the first stage and reduced form regression are the same as the event study design (Borusyak et al., 2021b). All of the estimates are similar to the main results. However, their effective F-statistics are less than the critical values that guarantee a relative bias of less than 10%. This implies that all of these estimates suffer from many-weak instruments problem (Table A5, A6).[22]

Finally, rather than six months as a threshold to define a program county, we use alternative thresholds, which are 0 and 12. The results are similar to the main result (Table A7, A8).

Validity of the identification assumption

In this section, we test the common trend assumption, key identification assumption, to establish the causal relationship in (5).

We implement a pre-trends test for NRPS enrollment and fertility to test this. If the common trend assumption holds in this case, the coefficients of pre-trends should not be

[21] This means that rather than using $Z_{c,t}$ defined in section 5.3, we use $1\{Dur\_year_{c,t} = l\}$ as IVs.

[22] In addition, rather than using the dummy variables, we use years of the duration and squared value of that as continuous IVs (not reported here). The estimates are similar to the main results.

significantly different from 0. Figure 3-5 show dynamic effects of NRPS implementation on the enrollment rate and fertility. All of the pre-trend coefficients are insignificant, and the p-value of the F-test is higher than 10% in all specifications.

Unfortunately, it is impossible to do pre-trends test for educational expenditures since expenditure data before 2009 is unavailable. Instead, we estimate the effect of NRPS implementation on the spending with urban samples. If unobserved characteristics between the program and non-program counties derived our results, the effects in urban areas would significantly differ from 0. Moreover, since the main target of this policy is rural areas, the effect in urban areas must be similar to 0. Thus, this estimation plays a placebo test that examines whether the effect exists or not when it "should not" exist. Table A9 presents the coefficient of NRPS implementation in urban areas. As expected, all estimates are not significantly different from 0, and the standard errors are substantial. This supports that our reduced form estimates in rural areas do not capture unobserved differences in these two groups.

Confounding factors

Finally, we estimate the effect of the NRPS on educational expenditure, controlling for additional confounding variables.

First, we consider several policies that may confound the NRPS in addition to our original control variables. In the 2000s, the Chinese government tried to reduce the economic gap between urban and rural areas. One attempt at this was educational and medical reform to improve rural areas' educational and medical levels. We consider two policies: Free Compulsory Education Reform and New Cooperative Medical Scheme, initiated in 2006 and 2003. These two policies provided tuition fees and miscellaneous fee exemptions for students and health insurance coverage for people in rural areas.[23] These

[23]See Tang et al. (2020) for more about Compulsory Education Reform and Liu (2016) for New Cooperative Medical Scheme.

policies may be why we estimate the large effects on educational spending. Table A10 shows estimates, adding the durations of educational and medical reforms as controls. After controlling for these two reforms, estimates do not change much. Thus, our estimates do not capture other reforms in rural areas the central government initiated in the 2000s.

Another concern is that our estimates capture the effect of the NRPS from working-age parents who pay a premium and grandparents who receive a basic pension. At the time the program is introduced, those older people aged 60 and over can directly receive the basic pension benefit without paying any premiums (Cheng et al. 2018, Yang and Chen 2022). In this case, grandparents who receive the basic pension would spend their pension incomes on their grandchild's education, which could be included in our estimates. To exclude the effect of receiving the basic pension, we estimate the effects of the NRPS with two samples. The first sample excludes families with any person who gets the basic pension. The other sample excludes families with any person eligible to receive the pension. In these two alternative samples, the estimates are similar or slightly larger than our main estimates, which implies that our main estimates do not include the effect of receiving the basic pension.[24] Finally, we estimate the effect of NRPS on financial support to non-coresident families, including support for parents and parents-in-law. The estimates are very small (82.39 RMB) compared to the estimate (577.27 RMB) in Panel B of Table 6, which is not statistically significant. This implies that the increase in the educational expenditure for children does not come from the reduction in financial support for parents.

[24] With the first sample, the effect of NRPS implementation is 621.08*** RMB (s.e.:223.17) while the estimate using the second sample is 888*** RMB (s.e.:171.94). These are similar to or larger than the main estimate (577.27 RMB) in Panel B of Table 6.

2.7 Conclusion

This paper considers how parents react to a pension program that can support them when they are old, which was initially the role of adult children in China. From the birth history and expenditure data in CFPS, we find that parents who enrolled in the NRPS spent more money on their children's education. However, we fail to see a significant effect on fertility from NRPS. When looking closely at fertility, they were more likely to give birth to a boy, whereas the probability of giving birth to a girl did not change. Moreover, we find parents increased educational expenditure for a son, but we fail to find a significant increase for a daughter. Looking at the cost in more detail, parents enrolled in the NRPS increased spending on extracurricular tutoring and transportation, which are easily manipulable.

This paper demonstrates that working-age parents do not prepare for their retirement at the cost of their children if the government introduces a formal pension scheme, which could be an alternative to adult children. Further, the results indicate that parents provide more support for their children after signing up for the pension scheme. This finding is consistent with previous literature that documented the improvements to the welfare of children by the NRPS. However, unlike previous papers that attributed the cause to a basic pension provided to the elderly, this paper shows the positive effects on children from working-age parents who pay the premium since the NPRS is a cheaper way to prepare for retirement. Finally, this paper suggests that we should not expect a pension scheme to weaken son preference, at least in the short term.

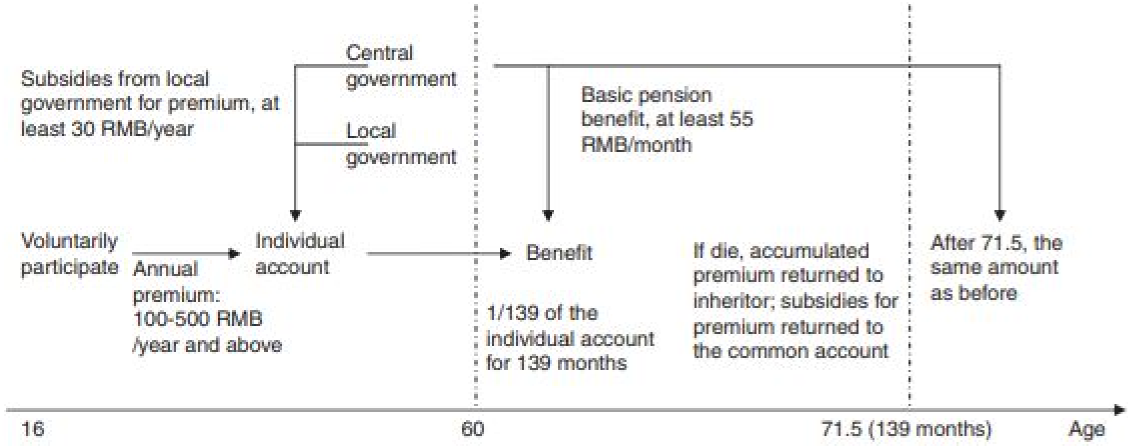
Similar to other papers on this subject, this paper has a limitation. The estimates contained in our study are the short-term effects of the pension program.[25] This is an important consideration because it is crucial to know about the long-term impacts of the NRPS for two reasons: a household cannot react to the policy by fertility over a short

[25]Since the Chinese government introduced the NRPS across all counties at the end of 2012, it is impossible to find counties as a control group to estimate long-term effects.

period, and, more importantly, it would take at least several years to change cultural behavior such as son preference, which was persistent in our case. Since a pension program would make sons less valuable, the NRPS may reduce son preference in the long run, like Bau (2021). Alternatively, parents may continue to prefer a son since China is a country with powerful son preference. Thus, it is crucial to determine any long-term effects to understand the impact of pension programs on cultural behaviors. We leave these for future research.

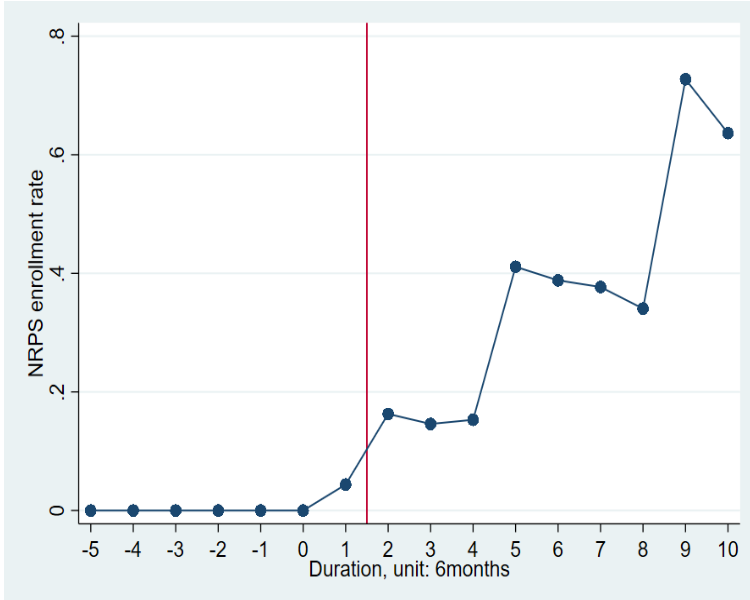
2.8 Figures and Tables

Figure 1. The new rural social pension program scheme



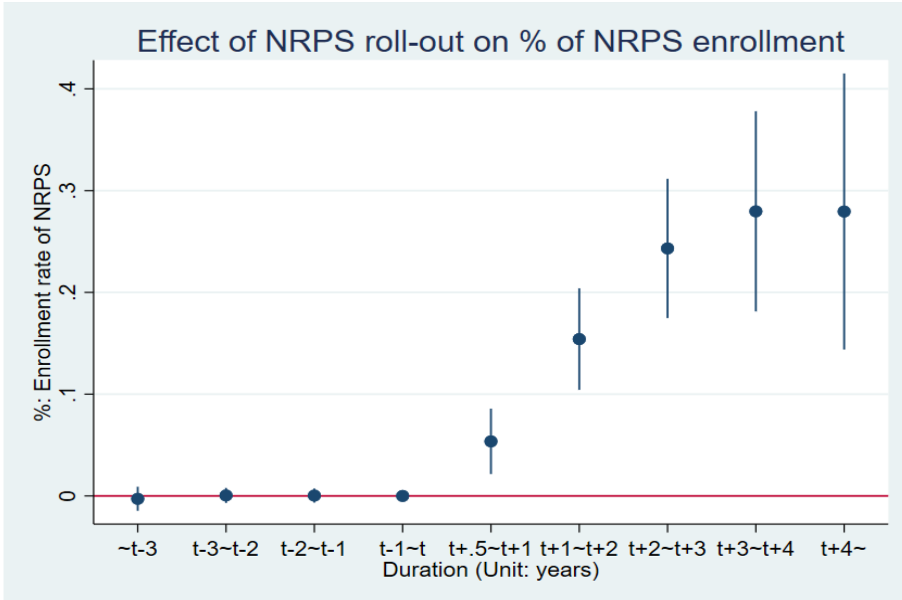
Source: Zhao et al. (2016)

Figure 2. The enrollment rate of NRPS over program duration



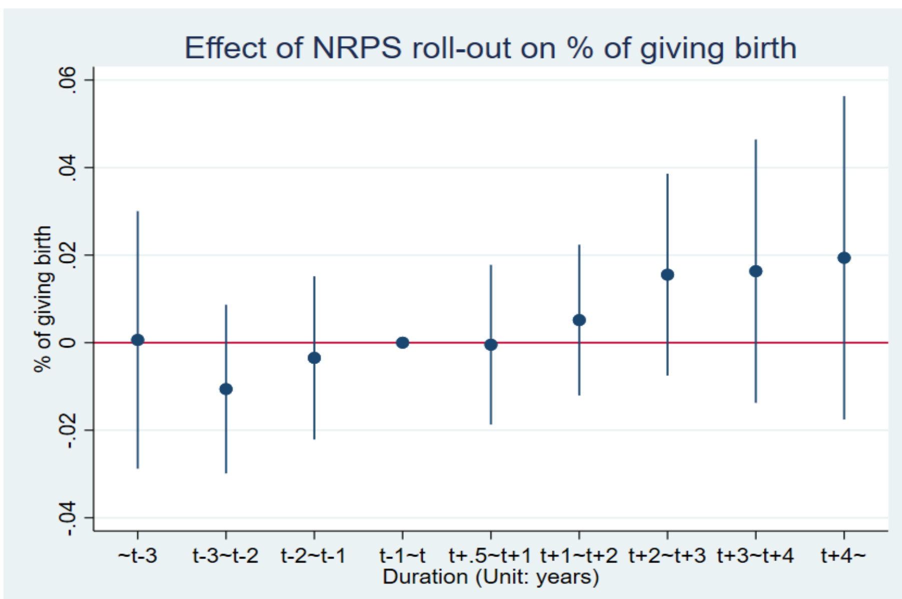
Note: Figure 2 shows the enrollment rate of NRPS among households with a child in this data over the duration of NRPS. Y-axis is the percentage of households with a child that signed up for NRPS. X-axis represents the duration of NRPS (time relative to NRPS roll-out) in counties. For instance, the enrollment rate of NRPS in counties in which the duration is in the first six month ($0 \leq Duration_{c,t} < 6$) is 4.4 percent (the first point on the left side of red vertical line) while the enrollment rate in counties in which the duration is between 6 and 12 months (the first point of right side of red vertical line) is 16.3%. Red vertical line splits counties by whether they are considered as program or non-program counties: Counties on the right side of vertical line ($Duration_{c,t} \geq 6$) are considered as program counties.

Figure 3. Dynamic effects of NRPS roll-out on NPRS enrollment



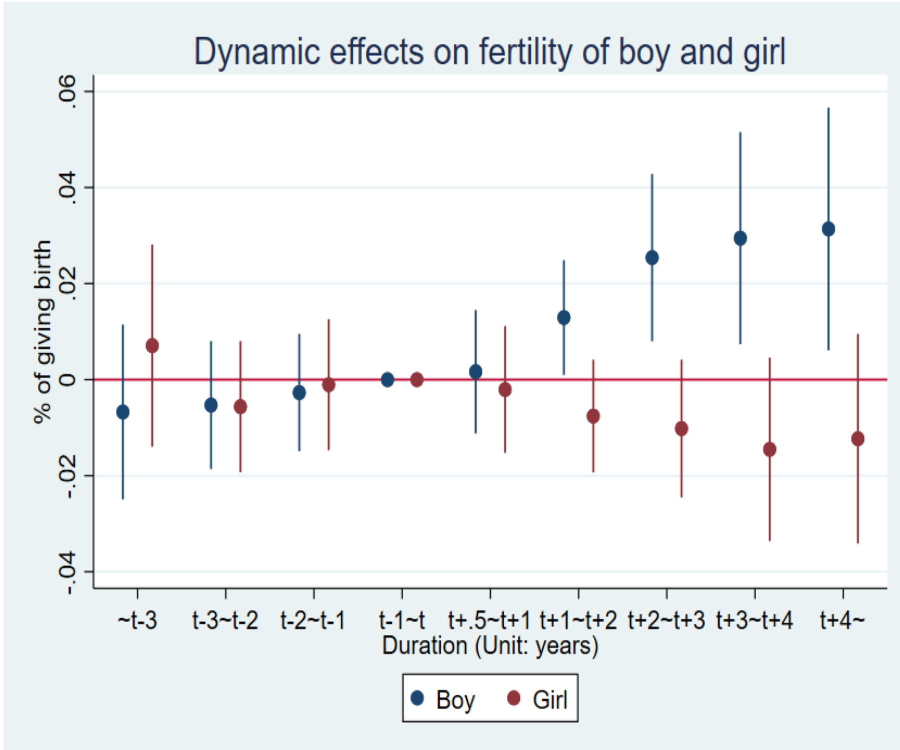
Note: Robust standard error clustered at county level in parentheses. Parents' age, their age squared, county and year fixed effects are controlled.

Figure 4. Dynamic effects of NRPS roll-out on fertility



Note: Robust standard error clustered at county level in parentheses. Parents' age, their age squared, county and year fixed effects are controlled.

Figure 5. Dynamic effects of NRPS roll-out on fertility of boy and girl



Note: Robust standard error clustered at county level in parentheses. Parents' age, their age squared, county and year fixed effects are controlled.

Table 1. The effect of the NRPS implementation on Q across π

	$\pi \in [0, p^{new})$	$\pi \in [p^{new}, p^{old}]$	$\pi \in (p^{old}, \bar{p}]$
Old-age security	No effect	$-Q(\pi)$	No effect
Altruism	$\lesseqgtr 0$	$\lesseqgtr 0$	$\lesseqgtr 0$
Altruism&Old-age security	No effect	$\lesseqgtr 0$	$\lesseqgtr 0$

Note: $\lesseqgtr 0$ means that the sign of the effect is ambiguous.

Table 2. Year of NRPS roll-out and the number of counties

Year of the NRPS roll-out	Number of counties
2008	70 (43%)
2009	30 (18%)
2010	26(16%)
2011	29 (18%)
2012	9 (5%)

Note: Following Xi et al. (2019), we calculate the earliest date in the data that enrollee paid premium for and received a basic pension from NRPS in a county and use the earliest date in a county as the date of the implementation in each county.

Table 3. Sample restrictions

	Fertility	Edu. expenditure
Available years	From 2004 to 2012	2009, 2011
Sample restriction		
Counties used in the analysis	All counties	Counties that implemented NRPS after June 2009
Hukou	Agricultural hukou	Agricultural hukou
Age of parents	$\leq$60 + mother$\leq$50	$\leq$60
Exclude outliers?	No	Yes, top 1% in the expenditure is excluded
Number of sample households	5,040	2,029
Number of sample counties	164	73

Table 4. Descriptive statistics of household characteristics

Variable	Mean	SD.	Mean	SD.	Mean	SD.
	Full Sample		09/10 Wave		11/12 Wave	
Educational expenditure	1228.55	1837.33	886.99	1492.85	1630.74	2104.93
NRPS enrollment	0.13	0.34	0.01	0.07	0.28	0.45
Duration of NRPS	6.89	9.49	0.59	1.27	14.30	9.63
Age of child	8.24	4.16	8.22	4.16	8.26	4.17
Age of father	36.06	7.04	35.82	6.83	36.33	7.28
Age of mother	34.02	6.97	33.84	6.78	34.22	7.18
Father's educational level	2.41	0.96	2.40	0.96	2.42	0.95
Mother's educational level	2.09	0.96	2.11	0.96	2.08	0.96
Amount of total assets	21.94	39.69	19.77	41.69	24.49	37.06
Number of kids	1.54	0.72	1.54	0.72	1.54	0.73
Sex ratio of kids	0.55	0.42	0.55	0.42	0.55	0.42

Note: Educational expenditure and total assets are adjusted to CPI in 2015. The unit of amount of total assets is 10,000 RMB.

Table 5. Effect of the NRPS on the indicator of giving birth to a baby

Dependent Variable	Indicator of giving birth to a baby			
Panel A	OLS	DID-IV		
NRPS Enrollment	-0.014**	0.051	0.053	0.050
	(0.006)	(0.060)	(0.060)	(0.060)
First Stage F-Statistic		33.42	33.79	34.64
Panel B		Reduced Form Regression		
Program Dummy		0.006	0.006	0.006
		(0.007)	(0.007)	(0.007)
County Fixed Effects	✓	✓	✓	
Family Fixed Effects				✓
Year Fixed Effects	✓	✓	✓	✓
Parents' Age, Age sqaured			✓	✓
Mean of Outcome	0.091	0.091	0.091	0.091
Observations	33,292	32,611	32,611	32,611

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Parents' age and age squared include father's and mother's age and their age squared. Column (1) in panel A indicates the effect of NPRS enrollment by OLS. Column (2)-(4) indicate the effects by using program dummy as IV. Column (2)-(4) in panel B includes the reduced form estimates.

Table 6. Effect of the NRPS on the average cost of education per child

Dependent Variable	Average Cost of Education			
Panel A	OLS	DID-IV		
NRPS Enrollment	227.56	1772.01***	1724.03**	1767.94**
	(165.79)	(640.36)	(720.35)	(756.41)
First Stage F-Statistics		49.70	51.70	48.11
Panel B		Reduced Form Regression		
Program Dummy		552.07***	577.27**	580.64***
		(176.46)	(222.14)	(254.19)
County Fixed Effects	✓	✓	✓	
Family Fixed Effects				✓
Year Fixed Effects	✓	✓	✓	✓
Household Characteristics			✓	✓
Mean of Outcome	1439.12	1241.49	1228.55	1298.00
Observations	2,993	2,850	2,208	1,600

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Household characteristics includes each parent's age, age suqred, educational level, the number of children, total amount of assets in a household, the number of generations in a family, and age of child. Column (1) in panel A indicates the effect of NPRS enrollment by OLS. Column (2)-(4) indicate the effects by using program dummy as IV. Column (2)-(4) in panel B includes the reduced form estimates.

Table 7. Effects of the NRPS on subcategories of educational expenditure

Dependent Varaible	Tuition	Book	Lodging	Extra. Tutoring	Transport.
Panel A			DID-IV		
NRPS Enrollment	101.41	47.11	-54.44	122.39	215.82**
	(230.27)	(112.68)	(81.60)	(84.06)	(92.73)
First Stage F-Statistic	48.11	48.11	48.11	47.89	48.11
Panel B			Reduced Form Regression		
Program Dummy	34.85	16.19	-18.71	42.35	74.16**
	(80.42)	(39.27)	(28.12)	(28.79)	(29.12)
NRPS Mean of Outcome	431.71	124.89	37.38	112.03	89.32
Observations	1,600	1,600	1,600	1,578	1,600

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Each parent's age, age suqred, educational level, the number of children, total amount of assets in a household, the number of generations in a family, age of child, family fixed effects, and wave fixed effects are controlled.

Table 8. Effects of the NRPS on the indicator of giving birth to boy and girl

Dependent Variable	Giving birth to a boy			Giving birth to a girl		
Panel A	DID-IV			DID-IV		
NRPS Enrollment	0.072*	0.072*	0.070*	-0.019	-0.018	-0.019
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
First Stage F-Statistics	33.42	33.79	34.64	33.42	33.79	34.64
Panel B	Reduced Form Regression			Reduced Form Regression		
Program Dummy	0.008*	0.008*	0.007*	-0.002	-0.002	-0.002
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
Mean of Outcome	0.048	0.048	0.048	0.044	0.044	0.044
Observations	32,611	32,611	32,611	32,611	32,611	32,611
County Fixed Effects	✓	✓		✓	✓	
Family Fixed Effects			✓			✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓
Parents' Age, Age sqaured		✓	✓		✓	✓

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Parents' age and age squared include father's and mother's age and their age squared.

Table 9. Effects of the NRPS on the total cost of education for first child by gender

	Total Cost of Education			
	Male	Female	Male	Female
Panel A	DID-IV			
NRPS Enrollment	2795.40***	1298.01	1937.54*	1846.97*
	(1059.11)	(832.78)	(1016.75)	(1039.04)
First Stage F-statistic	30.19	61.62	32.03	45.20
Panel B	Reduced Form Regression			
Program Dummy	885.62***	471.12	668.70*	639.98*
	(289.25)	(818.02)	(328.48)	(346.95)
Mean of Outcome	1286.10	1363.50	1379.10	1423.88
Observations	1,111	1,084	694	710
County Fixed Effects	✓			
Family Fixed Effects			✓	
Wave Fixed Effects	✓		✓	
Household Characteristics	✓		✓	

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Household characteristics includes each parent's age, age suqred, educational level, the number of children, total amount of assets in a household, the number of generations in a family, and age of child.

Chapter 3

Revisiting Unilateral Divorce Law and Divorce Rate: Using Dynamic Effect Estimators with Heterogeneous Treatment Effects

3.1 Introduction

The dynamics of family life have become very different since the 1950s. Nowadays, couples are more likely to get divorced, people opt to have less children, and people generally marry at an older age. Among these changes, social scientists have tended to focus on changes in the divorce rate, which rose from 2.2 per 1,000 people in 1960 to 5.0 per 1,000 people in 1985. Researchers have spotted a number of reasons for this change, including unilateral divorce law, which lowered the requirement for divorce from mutual to unilateral consent. In addition to social scientists' interest in the significant increase in the divorce rate, economists have focused on this law because the Coase theorem can be empirically tested by estimating the effect of this law.[1] However, despite many published empirical papers, the effect of unilateral divorce law on divorce and its dynamics remain contentious. Aside from the disputes between Peters (1986), Peters (1992), and Allen (1992) with cross-sectional data, there has still been no consensus even with panel data and with the data from different countries (Friedberg, 1998, Wolfers, 2006, Kneip et al., 2009, Lee and Solon, 2011). In particular, Friedberg (1998), Wolfers (2006), and Lee and Solon (2011) all used panel data from the United States but arrived at very different conclusions.

I revisit this discrepancy in prior estimates based on US panel data. I argue that the difference in such estimates indicates that the treatment effect of unilateral divorce law is heterogeneous across both states and time: The reason why the estimates in the three papers highlighted above are so different is that the treatment effects are heterogeneous while their empirical strategies assume homogeneous effects.[2] When the true effects are heterogeneous, the estimates derived from two-way fixed effects (TWFE) regression represent a weighted sum of the treatment effects. In the estimates, some of the weights are negative, which makes it impossible to arrive at a causal interpretation. In addition, in this case, the estimates via event study (ES) regression cannot be interpreted as the average treatment effect g periods after the initial treatment since they include treatment effects other than g periods after the initial treatment.[3] I provide alternative estimates

[1] Becker (1981) argues that from mutual consent to unilateral consent, the only change is the holder of the property right from a spouse who wants to remain married to who wants to divorce. Thus, this law must not affect the divorce rate. However, like many other applications to the Coase Theorem, Becker (1981)'s theoretical result depends on several strong assumptions. See Peters (1986), Chiappori et al. (2015).

[2] Heterogeneous treatment effects mean that $Y_{s,t}(1) - Y_{s,t}(0)$ is not restrictive where $Y_{s,t}(D)$ is a potential outcome and $D_{s,t}$ is a dummy variable indicating that state s has an unilateral divorce law in year t.

[3] Regarding more literature on this issue, see Abraham and Sun (2021), Borusyak et al. (2021a),

after gauging the magnitudes of the biases in the three papers.

Based on the alternative estimates, I find that unilateral divorce law decreased the divorce rate by 1.4%∼5.2%, much smaller than the effect estimated by Friedberg (1998), who found this law to increase the rate by 7.6%∼10.3% over 20 years. Regarding the dynamic pattern of the impact, I find that the divorce rate rose following the adoption of unilateral divorce law, although this rise was later reversed. This is slightly different from Wolfers (2006), who could not conclude the long-term effect of this change in the law.

The present paper summarizes recent findings in econometrics, especially concerning the heterogeneous treatment effects within repeated cross-section and panel data, and then applies those findings to the literature regarding the impact of unilateral divorce law on the divorce rate. This paper suggests comparing weighted and unweighted estimates when using the TWFE and ES estimation methods to detect heterogeneous treatment effects. Moreover, this paper proposes measures for gauging the bias associated with using ES estimation when the treatment effects are heterogeneous. These are consistent measures proposed by de Chaisemartin and D'Haultfuille (2020) about TWFE estimation. Finally, this paper warns applied economists who use TWFE estimates as the average treatment effects (e.g., Fadlon and Nielsen, 2019, Jaravel et al., 2018). This paper uses an example to demonstrate how TWFE estimates can differ from the average treatment effects. Thus, when the treatment effects are heterogeneous, it is more appropriate to calculate the average of the dynamic effects than to rely on TWFE estimates.

This paper proceeds as follows. In Section 2, I check identification assumptions for TWFE and ES, and if assuming heterogeneous effects from unilateral divorce law is reasonable. In Section 3, I replicate estimates of Wolfers (2006) and Lee and Solon (2011), who estimated dynamic effects, and see what econometric issue these estimates have and what drives differences between these two papers. And I provide alternative

Callaway and SantAnna (2021), de Chaisemartin and D'Haultfuille (2020), Goodman-Bacon (2021).

estimates. I do the same thing with Friedberg (1998) that estimated the effect by TWFE in Section 4. Section 5 concludes.

3.2 Identification Assumption and Heterogeneous Effects

In this section, I address an identification assumption and whether heterogeneous effects are plausible in the context of unilateral divorce law.

After Friedberg (1998), many studies on unilateral divorce law have exploited timing variations of unilateral divorce law across states, using two-way fixed effects regression (TWFE). TWFE regression refers to linear regression with treatment and time and group dummies, which is considered as a natural extension to difference in differences with 2 periods and 2 groups.

Regression 1. Two-way Fixed Effect Regression

$$Divorce_{s,t} = \alpha_s + \gamma_t + \beta_{fe} \cdot D_{s,t} + \epsilon_{s,t} \tag{3.1}$$

where $D_{s,t} = \mathbf{1}\{t \geq E_s\}$ *and* E_s *is the first year in which state* s *adopted the divorce law.*[4]

Event study (ES) regression divides a treatment dummy into several dummy variables by periods of exposure to treatment. Let $\mathcal{G}$ collect disjoint sets g of relative years between year t and the adoption year E_s. In the context of unilateral divorce law, Wolfers (2006) used the following regression equation:

[4]Between 1955-1988, there were ten years in which states adopted this law. I use these years as a notation for a group, that is, $e \in \{1969, 1970, \cdots, 1985, \infty\}$ where ∞ denotes a group that did not adopt the unilateral law over sample periods and plays the role of control group. See Table S1.

Regression 2. Event Study Regression

$$Divorce_{s,t} = \alpha_s + \gamma_t + \sum_{g=0}^{7} \beta_g \cdot D^g_{s,t} + \epsilon_{s,t} \tag{3.2}$$

where $D^0_{s,t} \equiv \mathbf{1}\{0 \leq t - E_s \leq 1\}$, $D^1_{s,t} \equiv \mathbf{1}\{2 \leq t - E_s \leq 3\}, \cdots$ *and* $D^7_{s,t} \equiv \mathbf{1}\{t - E_s \geq 14\}$.

TWFE and ES regression aim to estimate average treatment effects with appropriate interpretations for each parameter: we want to have average treatment effect from β_{fe} and average treatment effect g periods after the adoption from β_g. For β_{fe} and β_g to have causal interpretation, there is one necessary assumption to hold: common trend assumption.

Assumption 1. (common trend assumption)
For all integers $\ell \geq 1$, $\mathbb{E}\big(Y_{s,t}(0) - Y_{s,t-\ell}(0)\big|E_s = e\big)$ *does not vary across* e.

This assumption states that units in any group would have experienced the same trends in their mean outcomes if they had remained untreated. Under this assumption, we aim to get the average treatment effect on the treated by subtracting the trend of a control group from treatment groups. Note that this assumption assumes the same evolutions on unobserved counterfactuals across groups, so we cannot fully test this assumption. Rather, by comparing pre-trends between groups, this assumption can be tested indirectly.[5]

Figure 1 shows divorce rate changes across groups from 1955 to 1988. Before 1969, the first year of unilateral divorce law, the trends in divorce rate seemed very similar, which supports the common trend assumption.

[*Figure 1 is located near here.*]

[5]That is testing the hypothesis that for all integers $\ell \leq -1$, $\mathbb{E}\big(Y_{s,t}(0) - Y_{s,t-\ell}(0)\big|E_s = e\big) = 0$ for all e. This is called pre-trends or no anticipation assumption.

Nevertheless, the common trend assumption may be too strong to be plausible; therefore, many applied economists run TWFE and ES regressions with covariates. In this case, it is possible to relax this assumption by conditioning on covariates.

Assumption 2. (CONDITIONAL COMMON TREND ASSUMPTION)
For all integers $\ell \geq 1$, $\mathbb{E}\big(Y_{s,t}(0) - Y_{s,t-\ell}(0)\big|\mathbf{X}_s, E_s = e\big)$ *does not vary across* e.

This assumption assumes that although the unconditional mean of trends might differ across groups, the trends conditional on $\mathbf{X}_s$ are not different. Many applied economists parametrize this assumption by

$$\mathbb{E}[Y_{s,t}(0) - Y_{s,t-\ell}(0)|\mathbf{X}_s, E_s = g] = (X_{s,t} - X_{s,t-\ell})' \cdot \mu + \lambda_t. \tag{3.3}$$

And with the parametrization, TWFE and ES regression equations are the following:

Regression 3. (TWO-WAY FIXED EFFECT ESTIMATION WITH COVARIATES)

$$Divorce_{s,t} = \alpha_s + \gamma_t + \beta_{fe} \cdot D_{s,t} + X'_{s,t} \cdot \mu + \epsilon_{s,t} \tag{3.4}$$

where $X_{s,t}$ *is a vector of state specific time trends in Friedberg (1998) and Lee and Solon (2011).*

Regression 4. (EVENT STUDY ESTIMATION WITH COVARIATES)

$$Divorce_{s,t} = \alpha_s + \gamma_t + \sum_{g=0}^{7} \beta_g \cdot D^g_{s,t} + X_{s,t} \cdot' \mu + \epsilon_{s,t} \tag{3.5}$$

where $X_{s,t}$ *is a vector of state specific time trends in Wolfers (2006) and Lee and Solon (2011).*

Since the common trend assumption implies counterfactual trends are homogeneous across groups, one might think that the effect is more likely to be homogeneous across s.

However, the roll-out of the unilateral divorce law opens many channels for the characteristics of states to have interaction effects. For example, some states have a community property law that requires couples to split properties equally upon divorce. In these states, the unilateral law can cause higher divorce rates than states without the community law since a less powerful spouse can get more benefits by divorce.[6] Moreover, the effects would diminish over time rather than be constant since people in the states with unilateral divorce law become more careful in marriage to avoid divorce. This example predicts heterogeneous effects across states and time.

In the following two sections, I introduce propositions de Chaisemartin and D'Haultfuille (2020) and Abraham and Sun (2021) found about bias from TWFE and ES regression: a central problem of these two regressions is that only when treatment effects are homogeneous, the coefficients can be interpreted as a causal effect.

3.3 Event Study Regression

In this section, I revisit Wolfers (2006) and Lee and Solon (2011) to address bias in ES regression. They interpreted $\widehat{\beta}_g$ from (2) and (5) as the effect of the unilateral law on divorce rate g periods after the adoption. For (5), they controlled for linear state-specific trends in one specification and linear and quadratic trends in the other. The difference between Wolfers (2006) and Lee and Solon (2011) is that the former estimated $\widehat{\beta}_g$ weighted by state population while the latter did not weight. Figure 2 indicates $\widehat{\beta}_g$ from Wolfers (2006) and Lee and Solon (2011).

[*Figure 2 is located near here.*]

Wolfers (2006) found a significant effect on the divorce rate right after states adopt the divorce law, and this effect diminishes over time. Ten years after the adoption,

[6]For another example if there are more males in the marriage market in some states, women may tend to divorce since they are more likely to find another suitable husband than the current one.

out of three specifications, the effect decreases in the first two specifications while the effect gets larger again in the quadratic specification. That is why he concluded "it is hard to draw any stronger conclusion on long-term effect." In contrast, Lee and Solon (2011) pointed out that Wolfer's result is not robust to weighting. They could not find as uniform dynamic patterns as Wolfers (2006) across specifications. The difference in estimates with and without weighting has caused dissent on the effect of unilateral law on divorce rates.

Why does weighting cause a huge difference in estimates? Abraham and Sun (2021)'s decomposition on β_g helps to understand this discrepancy.

Proposition 1. Abraham and Sun (2021)'s decomposition
Let $\Delta_e(g) \equiv \mathbb{E}\big(Y_{s,e+\ell}(1) - Y_{s,e+\ell}(0) | E_s = e, \ell \in g\big)$. *Under common trend and pre-trends assumptions, in Regression 2 and 4,*

$$\beta_g = \sum_{e \neq \infty} w^g_{e,g} \cdot \Delta_e(g) + \sum_{g' \neq g \wedge g' \geq 0} \sum_{e \neq \infty} w^g_{e,g'} \cdot \Delta_e(g').$$

Moreover,

$$\sum_{e \neq \infty} w^g_{e,g} = 1 \textit{ and } \sum_{e \neq \infty} w^g_{e,g'} = 0 \textit{ for each } g'.^7$$

This proposition shows that β_g can be decomposed into two parts: 1) $\Delta_e(g)$ that should be in β_g to be interpreted as average treatment effect g periods after the initial treatment, and 2) $\Delta_g(g')$ that should not be in β_g. Thus, β_g is weighted sum of $\Delta_e(g)$ and $\Delta_e(g')$ for $g' \neq g$ with weights $w^g_{e,g}$ and $w^g_{e,g'}$ respectively.

The only difference between OLS and WLS is $w^g_{e,g}$ and $w^g_{e,g'}$. This decomposition gives two lessons to us: 1) only when $\Delta_e(g')$ does not vary across e for each g', that is $\Delta_e(g') = \Delta(g')$, $\widehat{\beta}_g$ from OLS and WLS must be the same. However, when the effects are heterogeneous, $\Delta_e(g') \neq \Delta(g')$ for any g', estimates of OLS and WLS must be pretty

[7] $w^g_{e,g'}$ is the regression coefficient on $D^g_{s,t}$ from regressing $D^{g'}_{s,t} \cdot \mathbf{1}\{E_s = e\}$ on all $D^g_{s,t}$ and state and year fixed effects (and a vector of covariates).

different. This is consistent with the finding of Solon et al. (2015) that the difference between estimates of OLS and WLS is evidence for unmodeled heterogeneity. 2) When treatment effects are heterogeneous, β_g cannot be interpreted as treatment effect g periods after the adoption since a) $w^g_{e,g}$ in the first term of β_g attached on the corresponding treatment effects can be negative and b) β_g includes treatment effects g' periods after the adoption other than g periods.

When we apply this decomposition in our case, Figure 2 is evidence for heterogeneity in effects. Since $\Delta_e(g')$ are different across e and g', estimates for the dynamic effects from Wolfers (2006) and Lee and Solon (2011) are pretty different. And because the effects are not homogeneous, the estimates in Wolfers (2006) and Lee and Solon (2011) cannot be interpreted as a weighted average of treatment effects g periods after the adoption. To gauge the bias in $\widehat{\beta}_g$, I calculate sum of absolute value of $w^g_{e,g'}$[8] and check if any $w^g_{e,g}$ is negative. The first measure's idea is that if the absolute value of each $w^g_{e,g'}$ is large, the effects that should not be included for the right interpretation affect β_g much. Thus, large $\sum_{g'\neq g \wedge g'\geq 0}\sum_{e\neq\infty}|w^g_{e,g'}|$ implies that β_g are affected largely by effects that should not be included.

[*Table 1 is located near here.*]

Table 1 shows the first measure in Wolfers (2006). Considering that for β_g to have the right interpretation, the sum of absolute weights for $\Delta_e(g)$ and $\Delta_e(g')$ for $g' \neq g$ must be equal to 1[9] and 0[10], $\widehat{\beta}_g$ includes a large amount of different treatment effects. Moreover, I find negative weights in $w^7_{e,7}$ in $\widehat{\beta}_7$ from quadratic trend specification, and the sum of them is -0.097.[11] Therefore, all of β_g in every specification are wrong estimates for treatment effects g periods after the adoption.

[8] Sum of $w^g_{e,g'}$ cannot be measure of bias since $\sum_{g'\neq g \wedge g'\geq 0}\sum_{e\neq\infty} w^g_{e,g'}$ is always 0.

[9] $\sum_{e\neq\infty} w^g_{e,\ell} = \sum_{e\neq\infty}|w^g_{e,g}| = 1$.

[10] $\sum_{g'\neq g \wedge g'\geq 0}\sum_{e\neq\infty} w^g_{e,g'} = \sum_{g'\neq g \wedge g'\geq 0}\sum_{e\neq\infty}|w^g_{e,g'}| = 0$.

[11] Except $\widehat{\beta}_7$ in the quadratic specification, I fail to find negative $w^g_{e,g}$.

To remove bias in each $\widehat{\beta}_g$, I use estimators proposed by Abraham and Sun (2021) (AS estimator) and de Chaisemartin and D'Haultfuille (2020) (DD estimator).[12] Both estimators are proposed to solve negative weights and exclude dynamic effects unrelated to $D^g_{s,t}$. However, their estimators are different from each other in two ways: First, DD estimator uses not-treated group and late-treated group as a control group, while AS estimator uses only not-treated group as a control group. Therefore, DD estimator uses more samples. Second, for inference, standard errors for the former estimator are calculated by bootstrap, whereas the latter uses an estimator for asymptotic variance. Therefore, these two estimators are a complement to each other, not a substitute.

Figure 3 shows alternative estimates to regression 2 and 4.

[*Figure 3 is located near here.*]

The patterns from new estimators are more similar to Wolfers (2006) than Lee and Solon (2011).[13] So I exclude dynamics from Lee and Solon (2011) in Figure 3. The divorce rate rose sharply in all specifications and estimators, and the effect gets smaller afterward except in the quadratic specification estimated by AS estimator.

There is a still spike 14 years after the adoption, which made Wolfers (2006) hard to conclude the long-term effects on the divorce rate. However, in this case, DD estimator estimates are preferred to AS estimator estimates in the quadratic specification. Note that the DD estimator uses more samples by using the late-treated group as the control group for the early-treated group. Thus, if the late-treated group is a good control group for the early group, then the DD estimator is more reliable since they use more samples.[14] The first evidence of this argument is Figure 1. Before 1969, trends of divorce

[12]See appendix for steps I follow to estimate alternative estimates.

[13]Since we are more interested in average treatment effects weighted by the number of population, I weight on the number of population in each state rather than $\frac{1}{N}$ where N is the number of states. See the appendix for steps I follow to estimate alternative estimates.

[14]I want to emphasize that using more samples is not always preferred: For instance, Dobkin et al.

rates were very similar across all groups, which supports that the late-treated group can be used as a control group for the early-treated group. Second, when I add these groups as control group, I fail to find any pre-trend and these estimates are very close to 0.[15] Thus, unilateral divorce law increases the divorce rate in the first ten years, but the effect diminishes over time.

Finally, considering estimates from the DD estimator are the most preferred, our bias measures gauge the magnitude of bias pretty well: compared to no-covariate and linear specifications, $\sum_{g'\neq g \wedge g'\geq 0}\sum_{e\neq\infty}|w^{g}_{e,g'}|$ in the quadratic specifications are larger in every $\widehat{\beta}_g$ and the difference between estimates of WLS and DD estimator in this specification are more significant than the other two specifications.

3.4 Two-way Fixed Effects Regression

In this section, I revisit Friedberg (1998) to address bias in TWFE regression. TWFE regression has been used to estimate the average treatment effect on treated when there are multiple groups and periods. Moreover, the estimate from TWFE regression has long been believed as the average of dynamic effects. Friedberg (1998) first estimated the impact of unilateral law on divorce rate with TWFE regression, and Wolfers (2006) estimated this with longer sample periods. And, again, Lee and Solon (2011) estimated without weighting on state population.

[*Table 2 is located near here.*]

(2018b) used people who had a hospital admission in the late wave (late-treated group) as a control group for those who had the admission in the early wave (early treated group) to estimate the impact of health shock. They did not use those who never had the admission (never-treated group) because they are more likely to be different from who had the admission at some point. For instance, they are likely to be more educated and wealthier, so they invest more in their health.

[15] See Figure S1-A.

Table 2 is the result of Friedberg (1998)[16] and Lee and Solon (2011). Here is the interpretation of Friedberg (1998): Her preferred estimates are estimates from the linear or quadratic specification. This is because state-specific trends help to control for state-specific time-variant unobserved factors, making groups more comparable. From estimates with the state-specific trends, she concluded that unilateral divorce law caused higher divorce rates significantly, which is a different finding from Peters (1992). However, her estimates are not robust to weighting: After removing weighting on state population, Lee and Solon (2011) failed to find any significant positive effect of this law.

The difference in OLS & WLS estimates is due to heterogeneity in effects. de Chaisemartin and D'Haultfuille (2020) decomposed β_{fe} into several causal effects:

Proposition 2. (de Chaisemartin and D'Haultfoeuille's decomposition) *Let* $\Delta_{e,t} \equiv \mathbb{E}\big(Y_{s,t}(1) - Y_{s,t}(0) | E_s = e\big)$*. Under Assumption 1, in Regression 1 and 3,*

$$\beta_{fe} = \sum_{(e,t):D_{s,t}=1} \frac{N_{e,t}}{N_1} \cdot w_{e,t} \cdot \Delta_{e,t}.$$

where $w_{e,t} = \frac{\epsilon_{e,t}}{\frac{1}{N_1}\sum_{(e,t):D_{s,t}=1} \epsilon_{e,t} N_{e,t}}$.[17] *Moreover,* $\sum_{(e,t):D_{s,t}=1} w_{e,t} = 1$ *and* $w_{e,t}$ *can be negative.*

This proposition explains the difference between estimates from OLS and WLS: If $\Delta_{e,t}$ is homogeneous across e and t, β_{fe} is just Δ from OLS and WLS. However, Table 2 refutes homogeneity. In addition, as seen in Section 3, our treatment effects are more likely to be heterogeneous across e and t. Proposition 2 also shows that when the treatment effects are heterogeneous, β_{fe} may not be interpreted as the weighted average of treatment effects

[16] I estimated with samples covering longer years than Friedberg (1998) and therefore, the estimates are same with Wolfers (2006) TWFE estimtates.

[17] $N_{e,t}$ and N_1 are the number of samples in group e in year t and that of treated samples. $\epsilon_{e,t}$ is residual of regression of $D_{s,t}$ on state and year fixed effects.

due to negative weights.[18] To gauge the bias, I calculate the sum of negative weights in each specification in Figure 4.[19]

[Figure 4 is located near here.]

Perhaps most interestingly, the estimates Friedberg (1998) preferred have much more negative weights. This implies that the significantly positive estimates may be due to these negative weights. In an extreme case, the sign of estimates could be different from the true effect. Also, including covariates may help common trend assumption to hold, but this may cause more bias in estimates by increasing negative weights. Figure 4 confirms this.

As alternative estimates, I calculate the average of dynamic effects from section 3 weighting on state population.[20] This estimate is closer to what we want from TWFE because 1) this estimate is calculated by averaging unbiased and consistent dynamic effects, and 2) this does not have negative weights. Table 3 compares these estimates with previous papers' results.

[Table 3 is located near here.]

All estimates using AS and DD estimators are negative except in the quadratic specifications with AS estimators. However, as estimates of DD estimator in this specification are preferred (see Section 3), $\widehat{\mathbb{E}}(\Delta_{e,t})$ from DD estimator is preferred to that of AS estimator. Even if $\widehat{\mathbb{E}}(\Delta_{e,t})$ from AS estimator is true, this is 37% lower than $\widehat{\beta}_{fe}$ in the previous paper and is not statistically significant anymore.

Second, our measure in Figure 4 predicts the magnitudes of bias in $\widehat{\beta}_{fe}$ well. Our estimates show that the unilateral divorce law decreased the divorce rate on average over

[18] Goodman-Bacon (2021) showed that β_{fe} is the weighted average of the treatment effects when the effects are constant over time but vary acorss units.

[19] I use `twowayfeweights` command in Stata. See de Chaisemartin and D'Haultfuille (2020).

[20] See appendix for steps I follow to estimate alternative estimates.

20 years, which are opposite conclusion from Friedberg (1998). Considering the average number of divorces per 1,000 people is 4.6, our estimates implied that this law decreased the divorce rate by from 1.4%$(= -\frac{0.065}{4.6} \times 100)$ to 5.2%$(= -\frac{0.242}{4.6} \times 100)$. The magnitude of effects estimated by alternative estimators are smaller, and the sign of them is opposite to estimates in Friedberg (1998).

3.5 Conclusion

TWFE and ES regression are widely used to estimate average treatment effects on treated. These methods have been believed to calculate the average because they look like a natural extension of difference in differences, which provides a causal effect in an obvious way. However, unlike this belief, recent econometric papers showed that estimates could be very far from actual treatment effects, especially if effects are heterogeneous.

I revisit previous papers estimating the effect of unilateral divorce law on the divorce rate with the above two methods. The effect has remained a puzzle because the estimated patterns with and without weighting on state population are so different. The decomposition of coefficients in TWFE and ES shows that this discrepancy would be evidence of heterogeneity in effects. When they are heterogeneous, TWFE and ES regression can no longer provide an average of treatment effects.

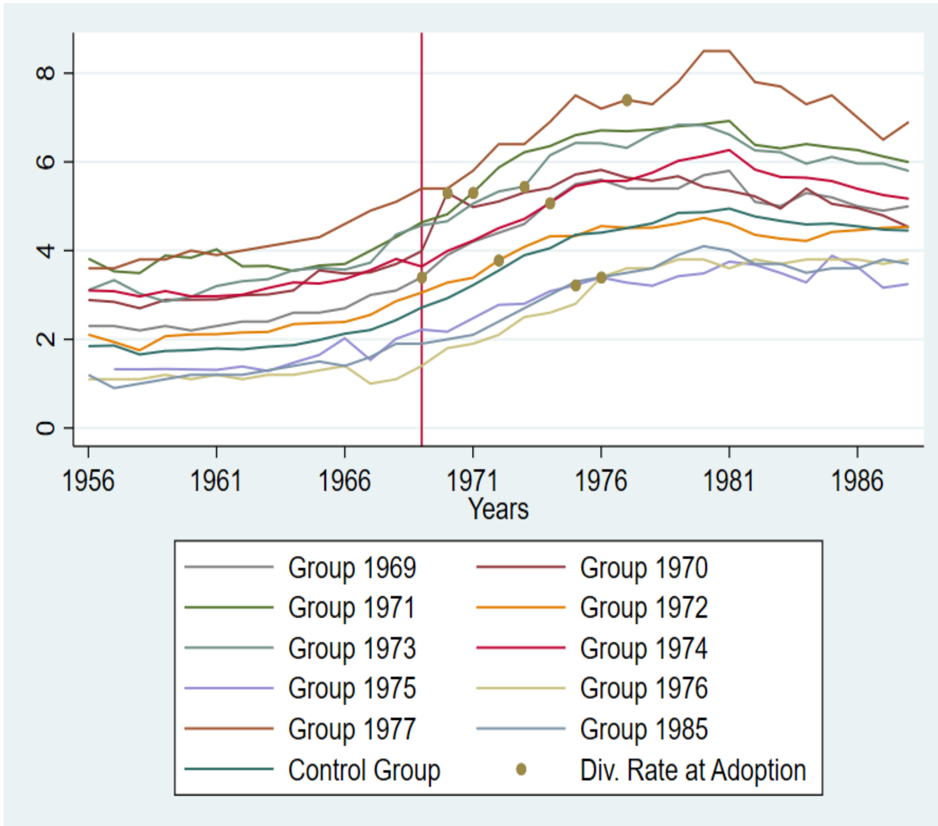
By using estimators robust to heterogeneous effects, I find that the divorce rate rose sharply right after states adopted unilateral divorce law, after which the effect decreased. Based on these dynamic effects, I re-calculate the average of dynamic effects, which is an alternative estimate for TWFE regression. This number shows that this divorce law affected divorce rates by a small amount on average, which is closer to Friedberg (1998)'s a least preferred estimate. So, this paper shows that ignoring heterogeneity in TWFE and ES can lead to the wrong conclusion.

This paper gives a lesson to applied economists who use these two methods: when

OLS and WLS estimates are very different, they may need to model heterogeneity (Solon et al., 2015) or 2) need to use alternative estimators robust to heterogeneity in effects.

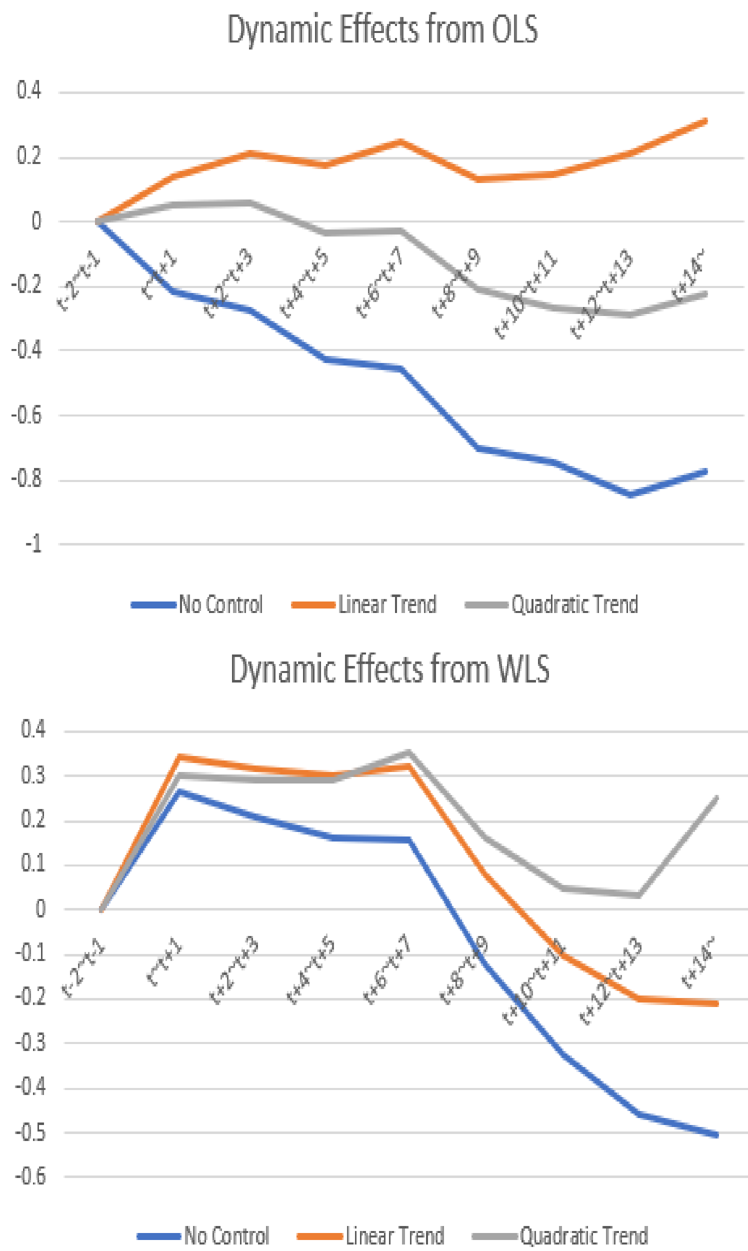
3.6 Figures and Tables

Figure 1: Trends of Divorce Rate across Groups over Sample Periods.



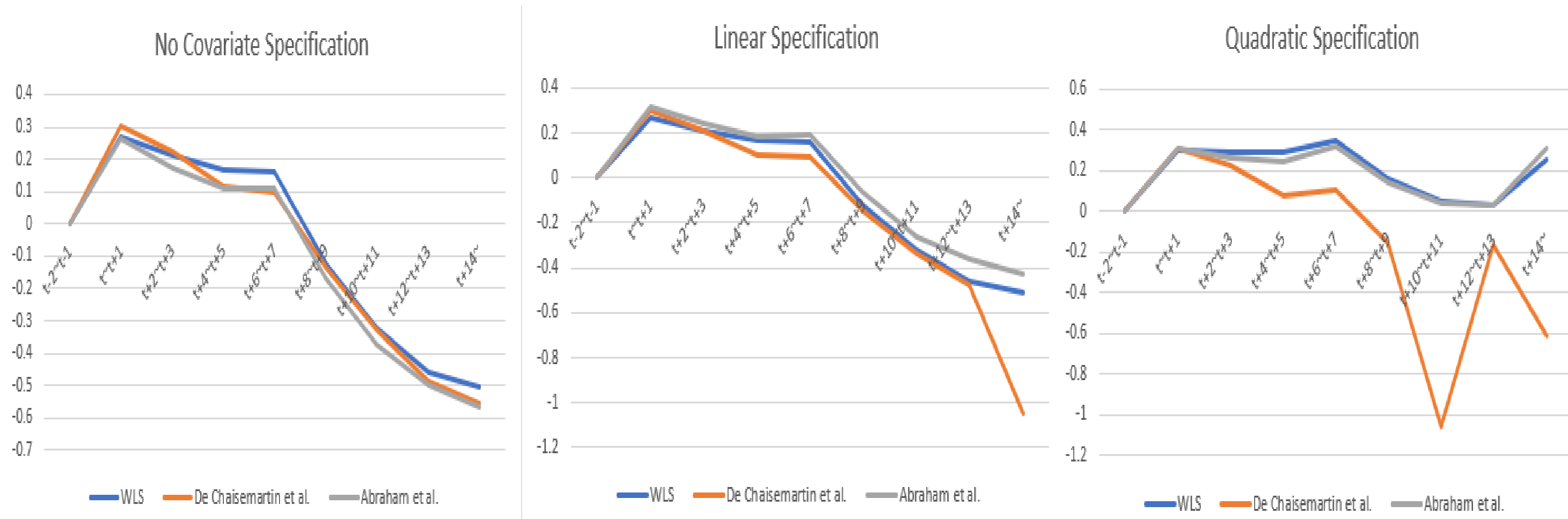
Notes: Figure 1 shows trends of the annual number of new divorces per thousand persons in different groups from 1955-1988. A group is defined by the year of the adoption of unilateral divorce law as Table S1. Green dots mean divorce rate in the year in which each group adopted this law.

Figure 2: Response of Divorce Rate to Unilateral Divorce Laws.



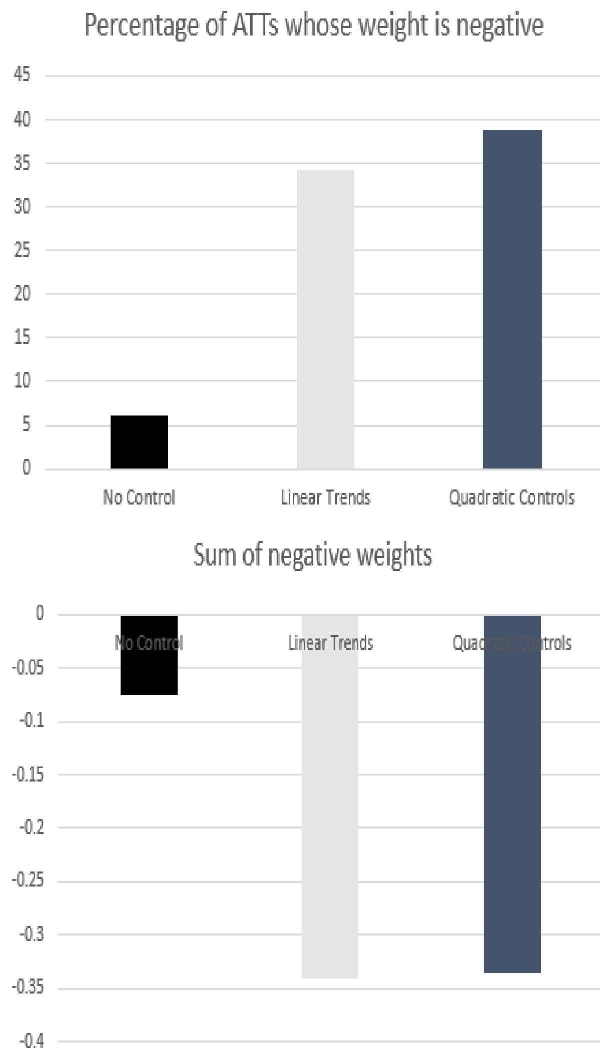
Notes: Figure 2 shows estimates of dynamic effects through event study estimation. The above figure shows estimates of Lee and Solon (2011) with weighting on state population whereas the below shows those of Wolfers (2006) without weighting. In each figures, there are three specifications, one does not control for any covariates, and the other two control for state-specific time trends.

Figure 3: Comparison between Wolfers (2006)(WLS) and alternative estimates.



Notes: WLS indicates estimates from Wolfers (2006). de Chaisemartin and D'Haultfuille (2020) and Abraham and Sun (2021). indicate estimates by using their alternative estimator. All regressions include controls for year and state fixed effects. Each standard error are clustered by state. For de Chaisemartin and D'Haultfuille (2020), standard errors are calculated by bootstrap 500 times.

Figure 4: Percent of ATTs with negative weights and sum of these.



Notes: Figure 4 indicates percentage of treatment effects whose weights are negative (left) and sum of negative weights (right) in β_{fe} in Friedberg (1998).

Table 1: Sum of absolute weights attached to different dynamic effects in each $\widehat{\beta}_g$

	No covariate	Linear-trend	Quadratic-trend
$\widehat{\beta}_0$	0.256	0.280	0.457
$\widehat{\beta}_1$	0.401	0.546	0.764
$\widehat{\beta}_2$	0.502	0.835	1.000
$\widehat{\beta}_3$	0.503	1.004	1.313
$\widehat{\beta}_4$	0.459	1.130	1.722
$\widehat{\beta}_5$	0.362	1.192	2.164
$\widehat{\beta}_6$	0.253	1.250	2.549
$\widehat{\beta}_7$	0.387	0.890	2.551

Notes: Table 1 shows sum of absolute value of weights $w^g_{e,g'}$ attached on $\Delta_e(g')$ across e and $g' \neq g$ in $\widehat{\beta}_g$ in Proposition 1. That is $\sum_{g' \neq g \wedge g' \geq 0} \sum_{e \neq \infty} |w^g_{e,g'}|$. Large values in $\sum_{g' \neq g \wedge g' \geq 0} \sum_{e \neq \infty} |w^g_{e,g'}|$ mean that the coefficient are affected largely by dynamic effects that should not be included in β_g.

Table 2: Response of Divorce Rate to Unilateral Divorce Law.

	Effect of unilateral divorce law		
$\widehat{\beta}_{fe}$ (WLS)	-0.055	0.476	0.334
	(0.151)	(0.202)	(0.151)
$\widehat{\beta}_{fe}$ (OLS)	-0.498	0.146	0.093
	(0.437)	(0.113)	(0.108)
Linear state-specific trends		✓	✓
Quadratic state-specific trends			✓

Notes: OLS indicates estimate without weighting on state population, which replicates Lee and Solon (2011). WLS indicates estimate with weighting on state population, which replicates Wolfers (2006). All regressions include controls for year and state fixed effects. Each standard error are clustered by state.

Table 3: Response of Divorce Rate to Unilateral Divorce Law.

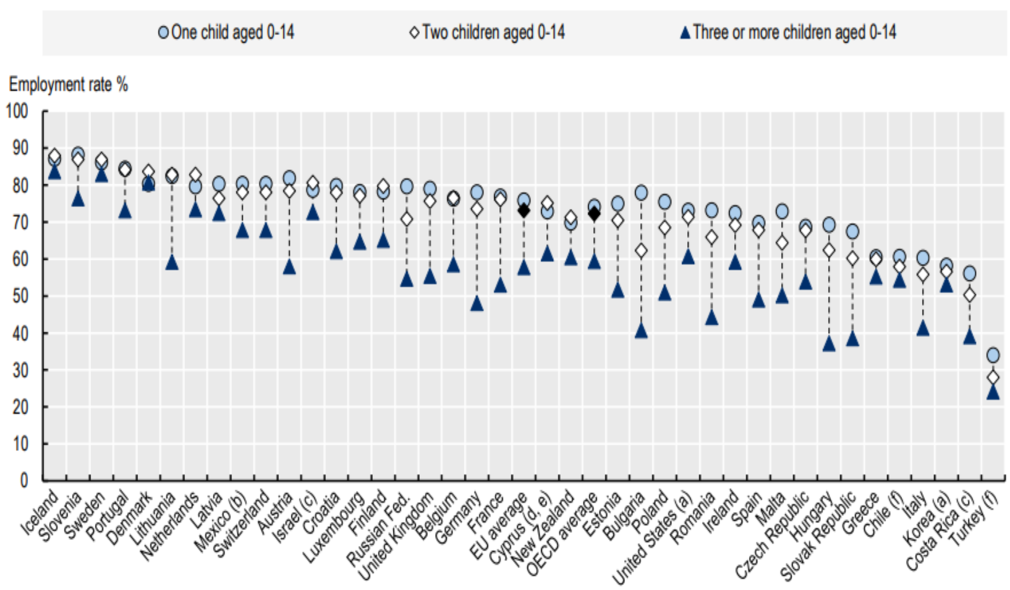
	Effect of unilateral divorce law		
$\widehat{\beta}_{fe}$	-0.498	0.476**	0.334**
	(0.437)	(0.202)	(0.151)
$\widehat{\mathbb{E}}(\Delta_{e,t})$ from AS	-0.165	-0.065	0.209
	(0.125)	(0.122)	(0.147)
$\widehat{\mathbb{E}}(\Delta_{e,t})$ from DD	-0.142	-0.242	-0.207
	(0.114)	(0.145)	(0.355)
Linear state-specific trends		✓	✓
Quadratic state-specific trends			✓

Notes: $\widehat{\beta}_{fe}$ indicates estimates from Wolfers (2006). $\widehat{\mathbb{E}}(\Delta_{e,t})$ from AS indicates estimates of average of dynamic effects estimated by Abraham and Sun (2021)'s alternative estimator. $\widehat{\mathbb{E}}(\Delta_{e,t})$ from DD indicates estimates of average of dynamic effects estimated by de Chaisemartin and D'Haultfuille (2020) estimator. All regressions include controls for year and state fixed effects. Each standard error are clustered by state. For $\widehat{\mathbb{E}}(\Delta_{e,t})$ from DD, standard errors are calculated by bootstrap 500 times.

Appendix A

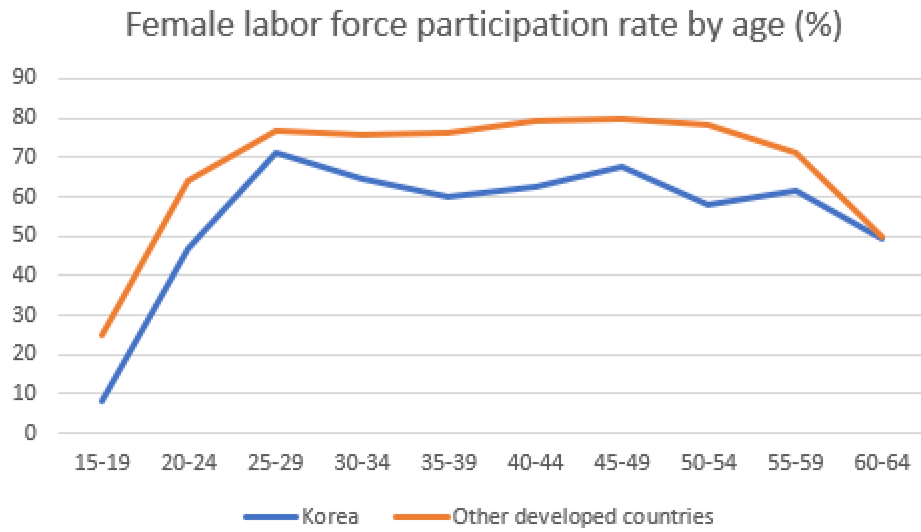
Appendix for Chapter 1

Figure A1: Maternal employment rates by number of children, 2019 (%)



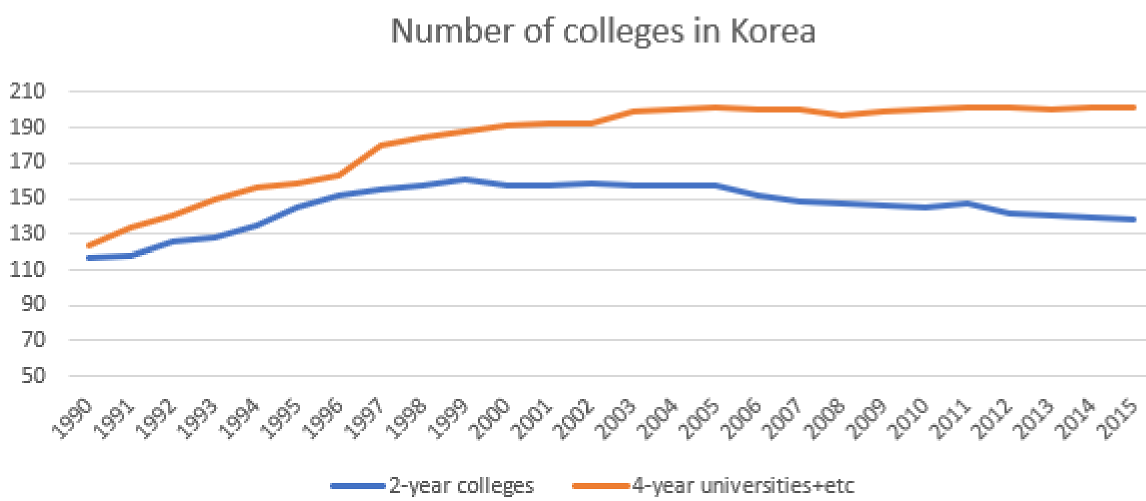
Note: Figure A1 shows maternal employment rate by the number of children in 2019 across countries. Source: OECD statistics

Figure A2: Female labor force participation rate by age in 2019 (%)



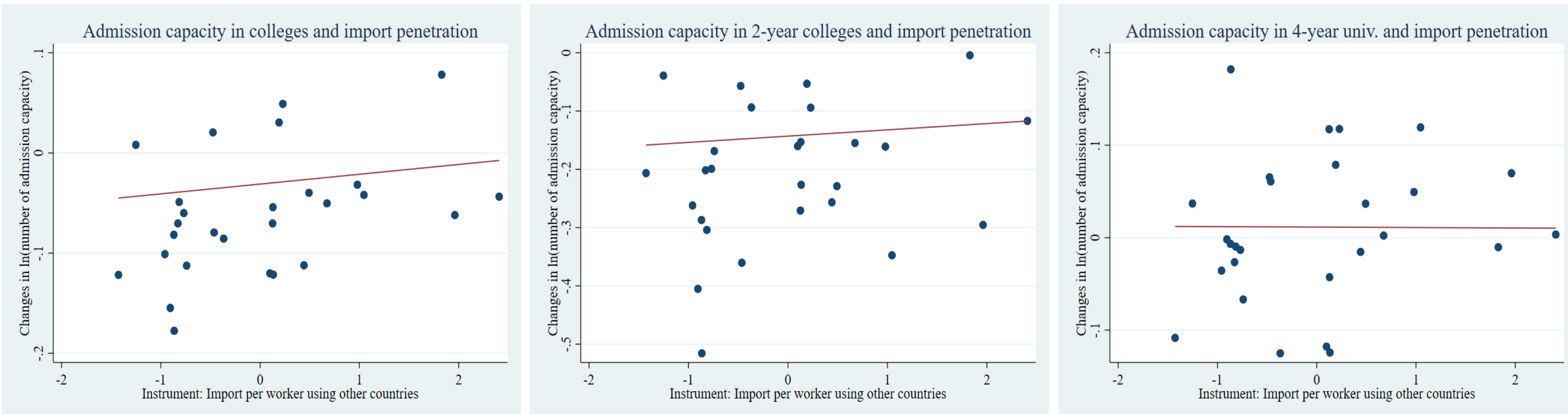
Note: Blue colored line represents female labor force participation rate in Korea. Orange colored line presents the average of the female rate in United Kingdom, Germany, France, Japan, and United States. Source: Korea Economic Research Institute.

Figure A3: Number of 2-year colleges and 4-year universities in Korea



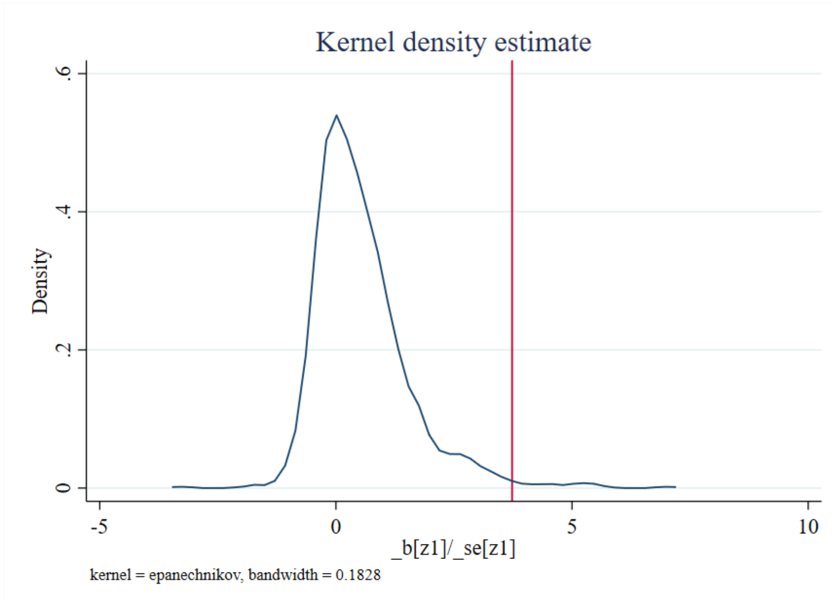
Note: Blue and orange colored lines present the number of 2-year colleges and 4-year universities in Korea between 1990 and 2015. Source: Korean Educational Development Institute

Figure A4: Number of 2-year colleges and 4-year universities in Korea



Note: The left figure shows a scatter-plot of all colleges' admission capacity and import penetration across provinces. The center and right figures show scatter-plots of 2-year and 4-year colleges' admission capacity and import penetration across provinces. Regression models are weighted by the population size in initial years of each period at the regional level. Source: Korean Educational Development Institute

Figure A5: Estimated coefficients from the permutation placebo tests (P-value: 0.012)



Note: I randomly assigned the import growth to industries and construct placebo instrumental variable. This figure plots the densities of the distribution of placebo instrumental variable effects from 1,000 random assignments. A Red line shows the estimated effect from the actual instrumental variable. A p-value is the proportion of placebo estimates that are equal to or larger in absolute value than the corresponding estimate from the actual data.

Table A1-a: Descriptive statistics by year: outcomes.

Per 100,000 residents	2000	2005	2010	2015
College enrollment (All)	1110.47	975.69	1017.35	866.08
2-yr college enrollment (All)	348.62	271.49	317.90	290.33
4-yr univ. enrollment (All)	761.85	704.20	699.45	575.74
College enrollment (Male)	596.51	517.32	523.33	431.24
2-yr college enrollment (Male)	188.46	140.56	158.17	141.41
4-yr univ. enrollment (Male)	408.04	376.76	365.15	289.83
College enrollment (Female)	513.96	458.36	494.02	434.83
2-yr college enrollment (Female)	160.15	130.92	159.73	148.92
4-yr univ. enrollment (Female)	353.81	327.43	334.29	285.91
N	50	50	50	50

Table A1-b: Descriptive statistics by year: Regional characteristics and trade.

Regional level	2000-05	2005-10	2010-15
% female	0.51	0.51	0.51
% 15 to 29	0.24	0.21	0.19
% Employed	0.40	0.41	0.47
% college degree	0.18	0.23	0.27
% manufacturing emp.	0.20	0.18	0.16
Import penetration using Korea imports	1.13	1.29	0.52
Import penetration using Other countries' imports	7.40	8.97	-0.83
N	50	50	50
Industry level			
Change in imports in Korea per one worker	28.23	54.09	-3.74
Change in other countries per one worker	160.45	194.29	-175.09
N	304	304	304

Table A2: Industry characteristics in 2000.

	# of workers per est.	Male ratio per est.
All industries	21.69	0.694
All importing industries	22.28	0.716
Importing industries in 50th PCTL	20.20	0.726
Importing industries in 70th PCTL	22.67	0.746
Importing industries in 90th PCTL	31.88	0.747

Note: The percentile is measured based on the growth of import values per worker in Korea from China between 2000 and 2015. As the percentile gets higher, that means the growth of imports in industries is higher. Source: Census on Establishments.

Table A3: Sensitivity analysis - Enrollment effects of local import penetration by specification

Panel A	Any CLG	2y CLG	4y UNIV	Any CLG-M	Any CLG-F	2y CLG-M	2y CLG-F	4y UNIV-M	4y UNIV-F
ΔIP_{it}^{ck}	196.7***	103.9***	92.81***	89.69***	107.0***	38.52**	65.38***	51.17***	41.64***
	(39.94)	(34.42)	(25.83)	(19.89)	(21.64)	(16.32)	(18.97)	(15.22)	(12.14)
	[39.13]	[32.01]	[23.04]	[16.09]	[26.41]	[14.22]	[19.40]	[12.34]	[14.22]

Included controls: Period fixed effects, lagged employment share

1st stage F-stat (Regional level): 32.48 / 1st stage F-stat (Industry level): 24.70

Panel B	Any CLG	2y CLG	4y UNIV	Any CLG-M	Any CLG-F	2y CLG-M	2y CLG-F	4y UNIV-M	4y UNIV-F
ΔIP_{it}^{ck}	111.2**	66.01**	45.16	56.98**	54.19**	22.82*	43.19***	34.16**	11.00
	(44.90)	(26.67)	(28.57)	(22.54)	(23.01)	(13.43)	(15.37)	(15.48)	(15.41)
	[48.14]	[28.29]	[25.28]	[23.76]	[24.79]	[14.92]	[14.76]	[12.34]	[14.44]

Included controls: Province-specific period fixed effects, period-specific lagged employment share

1st stage F-stat (Regional level): 12.95 / 1st stage F-stat (Industry level): 13.54

Panel C	Any CLG	2y CLG	4y UNIV	Any CLG-M	Any CLG-F	2y CLG-M	2y CLG-F	4y UNIV-M	4y UNIV-F
ΔIP^{ck}_{it}	106.8***	61.18***	45.65	55.14***	51.69**	21.34*	39.85***	33.81**	11.85
	(40.61)	(23.39)	(28.41)	(20.69)	(20.60)	(11.26)	(14.34)	(16.10)	(14.32)
	[43.11]	[23.57]	[25.38]	[21.71]	[21.83]	[11.38]	[13.49]	[13.26]	[13.37]

Included controls: Province-specific period fixed effects, period-specific lagged employment share, initial period of female share, 15-29 share

1st stage F-stat (Regional level): 11.75 / 1st stage F-stat (Industry level): 12.32

Standard errors clustered at the regional level are in parentheses while those clustered at the industry level in brackets.

Number of regions: 50 / Number of industries: 304.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ based on standard errors clustered at the regional level

Note: The coefficients represent the effect of one standard deviation increase in ΔIP^{ck}_{it} on outcomes. All observations are weighted by the population size in initial year of each period at the regional level.

CLG: College. 2y CLG: 2-year college. 4y. UNIV: 4-year University. M: Male. F: Female.

Table A4: Correlation between pre-determined covariates and imports of other countries

	Import growth of other countries
log sale	696.1
	(2412.0)
log workers	54.94
	(89.80)
Share of female	-33.07
	(207.4)
log production	-682.2
	(2427.6)
log value-added	13.28
	(49.26)
log wage	-152.2
	(101.3)
N	912
P-value of joint significance test	0.2873

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table reports estimates of regressions of the industry-level import growth in other countries IP^{co}_{jt} during 2000-2015 on a set of pre-determined industry characteristics. Industry characteristics are from Mining and Manufacturing Survey in Korea in 1999, 2004, and 2009. Numbers in parentheses are standard errors are clustered at the industry level. The P-value of joint significance test is p-value of the joint test that all coefficients above are equal to 0.

Table A5: Enrollment effects of local import penetration by the choice of countries

Panel A	Any CLG	2y CLG	4y UNIV	Any CLG-M	Any CLG-F	2y CLG-M	2y CLG-F	4y UNIV-M	4y UNIV-F
ΔIP_{it}^{ck}	106.8***	61.18***	45.65	55.14***	51.69**	21.34*	39.85***	33.81**	11.85
	(40.61)	(23.39)	(28.41)	(20.69)	(20.60)	(11.26)	(14.34)	(16.10)	(14.32)

1st stage F-statistic (Region): 11.75

Included countries: Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, Swiss, Canada, Singapore, United Kingdom, Norway, Sweden

Panel B	Any CLG	2y CLG	4y UNIV	Any CLG-M	Any CLG-F	2y CLG-M	2y CLG-F	4y UNIV-M	4y UNIV-F
ΔIP_{it}^{ck}	142.9**	99.11***	43.80	69.89**	73.02**	32.65**	66.46***	37.24*	6.556
	(63.75)	(32.49)	(38.38)	(31.11)	(33.23)	(15.34)	(19.85)	(20.41)	(20.39)

1st stage F-statistic (Region): 6.55

Among the countries from Panel A, top 10 countries trading with Korea are excluded, which are Australia, Germany, and Japan.

Panel C	Any CLG	2y CLG	4y UNIV	Any CLG-M	Any CLG-F	2y CLG-M	2y CLG-F	4y UNIV-M	4y UNIV-F
ΔIP_{it}^{ck}	145.4**	86.96**	58.44	72.99**	72.41**	30.93*	56.03***	42.06*	16.39
	(70.22)	(37.65)	(45.48)	(36.53)	(34.51)	(18.78)	(21.44)	(25.54)	(22.63)

1st stage F-statistic (Region): 5.76

Among the countries from Panel A, Asian countries are excluded, which are Japan and Singapore.

Note: Numbers in parentheses are standard errors are clustered at the regional level. The coefficients represent the effect of one standard deviation increase in ΔIP_{it}^{ck} on outcomes. All observations are weighted by the population size in initial year of each period at the regional level. Control variables are province-specific period fixed effects, period-specific lagged manufacturing employment shares, the share of females, and the share of people aged between 15 and 29 in the initial year of each period. CLG: College. 2y CLG: 2-year college. 4y. UNIV: 4-year University. M: Male. F: Female

Table A6: Bias of using college enrollments at local colleges as an outcome

	All students' enrollment	Local students' enrollments	All students' enrollment	Local students' enrollments
ΔIP^{co}_{it}	27.66*	30.87*	70.37***	57.41***
	(16.03)	(16.41)	(21.12)	(13.00)
Province-specific period fixed effects	X			
Period fixed effects			X	

N: 150.

Standard errors in parentheses

$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$

Note: Numbers in parentheses are standard errors are clustered at the regional level. The coefficients represent the effect of one standard deviation increase in ΔIP^{co}_{it} on outcomes. All observations are weighted by the sampling weight of the survey. All students' enrollments are local and non-local students' enrollments at local colleges. Local students' enrollments are college enrollments of students who still lives in regions in which they lived 5 years ago (local students). Control variables are period-specific lagged manufacturing employment shares, the share of females, and the share of people aged between 15 and 29 in the initial year of each period.

Table A7: Effects of local import penetration on population size

	ln(Total)	ln(Male)	ln(Female)
ΔIP_{it}^{co}	-0.00922	-0.00959	-0.00884
	(0.00858)	(0.00885)	(0.00831)
N		150	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Numbers in parentheses are standard errors are clustered at the regional level. The coefficients represent the effect of one standard deviation increase in ΔIP_{it}^{co} on outcomes. All observations are weighted by the population size in initial year of each period at the regional level. Control variables are province-specific period fixed effects, period-specific lagged manufacturing employment shares, the share of females, and the share of people aged between 15 and 29 in the initial year of each period.

Table A8: Balance check - regional characteristics in initial year of each period

	Female	% 15-29	% Working	% Working-femlae	% Above college	% Above college-female
ΔIP^{co}_{it}	-0.000384	0.00170	-0.00816*	-0.00312	0.00213	0.00124
	(0.00100)	(0.00344)	(0.00409)	(0.00205)	(0.00484)	(0.00228)
N	150 = 50 (Number of regions) × 3 (Number of periods)					

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Numbers in parentheses are standard errors are clustered at the regional level. The coefficients represent the effect of one standard deviation increase in ΔIP^{co}_{it} on outcomes. All observations are weighted by the population size in initial year of each period at the regional level. Control variables are province-specific period fixed effects, period-specific lagged manufacturing employment shares.

Table A9: Balance check - outcomes in initial year of each period

	Any CLG	2y CLG	4y UNIV	Any CLG-M	Any CLG-F	2y CLG-M	2y CLG-F	4y UNIV-M	4y UNIV-F
ΔIP^{co}_{it}	-13.76	-3.677	-10.09	-4.066	-0.309	-3.757	-9.699	-3.368	-6.331
	(13.42)	(6.329)	(7.600)	(6.847)	(3.248)	(4.160)	(13.65)	(6.974)	(7.480)
N	150 = 50 (Number of regions) $\times$ 3 (Number of periods)								

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Numbers in parentheses are standard errors are clustered at the regional level. The coefficients represent the effect of one standard deviation increase in ΔIP^{co}_{it} on outcomes. All observations are weighted by the population size in initial year of each period at the regional level. Control variables are province-specific period fixed effects, period-specific lagged manufacturing employment shares.

CLG: College. 2y CLG: 2-year college. 4y. UNIV: 4-year University. M: Male. F: Female.

Appendix B

Appendix for Chapter 2

B.1 Description of Independent Variables

- Age of father: Father's age in household i
- Age of mother: Mother's age in household i
- Age of child: Average of children in household i
- Education level of father: Categorical variable (1-7) of father's education in family i
- Education level of mother: Categorical variable (1-7) of mother's education in family i
- Number of generations: Number of generations in a family i
- Total amount of assets: Total amount of family i assets adjusted in 2015 CPI.
- Duration of Free Compulsory Education Reform: Duration of Free compulsory Education Reform in county c
- Duration of New Cooperative Medical Scheme: Duration of New Cooperative Medical Scheme in county c

B.2 Additional Tables

Table A1. Descriptive statistics of households with a child

	Without the NRPS	With the NRPS	Difference
Age of mother	32.92	34.1	-1.18***
	(0.099)	(0.200)	(0.222)
Age of father	34.81	35.88	-1.07***
	(0.099)	(0.198)	(0.222)
Number of children	1.48	1.47	-0.002
	(0.009)	(0.018)	(0.020)
Sex ratio of children	0.550	0.535	0.015
	(0.006)	(0.012)	(0.014)
Parents' Income	11024	10299	725.56
	(197.90)	(530.85)	(566.54)
Education level of mother	2.23	2.25	-0.027
	(0.014)	(0.029)	(0.032)
Education level of father	2.49	2.58	-0.095*
	(0.015)	(0.032)	(0.035)
Amount of total assets	25.97	24.73	1.230
	(1.230)	(1.125)	(1.666)

Note: t-test was applied for pairwise comparison. $***p < 0.01; **p < 0.05; *p < 0.1$. The numbers in parenthesis are standard errors. The unit of the amount of total assets is 10,000 RMB.

Table A2. A sample restriction in the expenditure analysis

Available year in data / Date that counties implemented NRPS	2009	2011
2008.1∼2009.6	Treated	Treated
2009.7∼2011.6	Untreated	Treated
2011.7∼2012.8	Untreated	Untreated

Note: Since the empirical strategy is basically difference in differences, it needs untreated and treated data before and after the roll-out of NRPS respectively in program counties (treatment group) while only untreated data is needed for non-program counties (control group). Therefore, data of counties that implemented NRPS before June of 2009 is excluded in the expenditure analysis.

Table A3. Estimation for the NRPS enrollment decision - First stage regression

Dependent Variable	NRPS enrollment		
	First Stage Regression		
Program Dummy	0.12***	0.12***	0.12***
	(0.02)	(0.02)	(0.02)
First Stage F-Statistics	33.42	33.42	33.47
County Fixed Effects	✓	✓	
Family Fixed Effects			✓
Year Fixed Effects	✓	✓	✓
Parents' Age, Age sqaured		✓	✓
Observations	32,611	32,611	32,611

Note: Table 5 presents estimates of the first stage regression when the main dependent variable is an indicator of giving birth to a baby. Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Parents' age and age squared include father's and mother's age and their age squared.

Table A4. Estimation for the NRPS enrollment decision - First stage regression

Dependent Variable	NRPS enrollment		
	First Stage Regression		
Program Dummy	0.31^{***}	0.33^{***}	0.34^{***}
	(0.04)	(0.05)	(0.05)
First Stage F-Statistics	49.70	51.70	48.11
County Fixed Effects	✓	✓	
Family Fixed Effects			✓
Year Fixed Effects	✓	✓	✓
Parents' Age, Age sqaured		✓	✓
Observations	2,845	2,208	1,600

Note: Table 6 presents estimates of the first stage regression where the main dependent variable is the average cost of education per a child in a household. Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Household characteristics includes each parent's age, age suqred, educational level, the number of children, total amount of assets in a household, the number of generations in a family, and age of child.

Table A5. Alternative estimates: Effect of the NRPS on the indicator of giving birth Alternative instrumental variables: 5 dummy variables indicating the duration in years.

Dependent Variable	Indicator of giving birth to a baby		
	Event Study-IV		
NRPS Enrollment	0.072^{*}	0.072^{*}	0.067^{*}
	(0.041)	(0.041)	(0.039)
First Stage F-Statistic	12.99	13.04	13.78
Observations	32,611	32,611	32,611
Mean of Outcome	0.091	0.091	0.091
County Fixed Effects	✓	✓	
Family Fixed Effects			✓
Year Fixed Effects	✓	✓	✓
Parents' Age, Age sqaured		✓	✓

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Parents' age and age squared include father's and mother's age and their age squared. Instruments are 5 dummy variables indicating whether a county had been exposed more than 6 months but less than 1 year, more than 1 year but less than 2 years, more than 2 years less than 3 years, more than 3 years but less than 4 years, and more than 4 years.

Table A6. Alternative estimates: Effect of the NRPS on the average cost of education

Alternative instrumental variables: 3 dummy variables indicating the duration in years

Dependent Variable	Average Cost of Education		
	Event Study-IV		
NRPS Enrollment	1268.43**	1228.55*	1652.33**
	(608.77)	(668.66)	(719.96)
First Stage F-Statistic	8.526	9.30	8.473
County Fixed Effects	✓	✓	
Family Fixed Effects			✓
Year Fixed Effects	✓	✓	✓
Household Characteristics		✓	✓
Mean of Outcome	1241.49	1228.55	1298.00
Observations	2,850	2,208	1,600

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Household characteristics includes each parent's age, age suqred, educational level, the number of children, total amount of assets in a household, the number of generations in a family, and age of child. Instruments are 3 dummy variables indicating whether a county had been exposed more than 6 months but less than 1 year, more than 1 year but less than 2 years, and more than 2 years.

Table A7. Alternative estimates: Effect of the NRPS on the indicator of giving birth

Alternative threshold to define program counties: 0 and 12 months

Dependent Variable	Indicator of giving birth to a baby		
Panel A	DID-IV. IV: 0≤Duration		
NRPS Enrollment	0.047	0.048	0.042
	(0.090)	(0.089)	(0.088)
First Stage F-Statistics	24.94	25.11	25.98
Panel B	DID-IV. IV: 12≤Duration		
NRPS Enrollment	0.063	0.063	0.062
	(0.044)	(0.044)	(0.043)
First Stage F-Statistics	42.50	42.78	43.49
Mean of Outcome	0.091	0.091	0.091
Observations	32,611	32,611	32,611
County Fixed Effects	✓	✓	
Family Fixed Effects			✓
Year Fixed Effects	✓	✓	✓
Parent's Age, Age squared		✓	✓

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively.

Table A8. Alternative estimates: Effect of the NRPS on the average cost of education

Alternative threshold to define program counties: 0 and 12 months

Dependent Variable	Average Cost of Education		
Panel A	DID-IV. IV: 0≤Duration		
NRPS Enrollment	1892.04**	1812.37*	1580.69
	(840.58)	(966.00)	(1029.87)
First Stage F-Statistics	46.94	45.26	45.09
Mean of Outcome	1246.46	1248.85	1334.07
Observations	2,072 (506)	1,638 (434)	1,200 (340)
Panel B	DID-IV. IV: 12≤Duration		
NRPS Enrollment	1380.12**	1581.29**	2202.90**
	(733.88)	(754.98)	(885.59)
First Stage F-Statistics	20.45	22.27	16.36
Mean of Outcome	1280.79	1271.72	1345.12
Observations	3,196 (1,137)	2,445 (910)	1,764 (668)
County Fixed Effects	✓	✓	
Family Fixed Effects			✓
Year Fixed Effects	✓	✓	✓
Household Characteristics		✓	✓

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Household characteristics includes each parent's age, age suqred, educational level, the number of children, total amount of assets in a household, the number of generations in a family, and age of child. The numbers in parenthesis in Observation row is the number of observations in the control group.

Table A9. Effects of NRPS on the average cost of education for each child in urban area

Dependent Variable	Average Cost of Education		
Urban Sample	Reduced Form Regression		
Program Dummy	853.66	932.51	-627.48
	(958.23)	(1131.84)	(1351.30)
Mean of Outcome	3556.08	3578.19	3578.19
Observations	668	564	564
County Fixed Effects	✓	✓	
Family Fixed Effects			✓
Wave Fixed Effects	✓	✓	✓
Household Characteristics		✓	✓

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Household characteristics includes each parent's age, age squared, educational level, the number of children, total amount of assets in a household, the number of generations in a family, and age of child. The table includes estimates of program dummy on the average cost of education per a child in urban area.

Table A10. Effects of the NRPS after controlling for confounding policies

Dependent Variable	Average Cost of Education		
Panel A	DID-IV		
NRPS Enrollment	1651.79*	1820.72**	1916.09**
	(791.06)	(786.21)	(782.10)
First Stage F-Statistics	47.16	47.39	43.37
Panel B	Reduced Form Regression		
Program Dummy	532.50**	597.18**	639.08**
	(240.24)	(239.76)	(246.85)
County Fixed Effects	✓	✓	
Family Fixed Effects			✓
Year Fixed Effects	✓	✓	✓
Household Characteristics		✓	✓
Confounding Policies	✓	✓	✓
Mean of Outcome	1198.98	1210.36	1277.45
Observations	2,041	1,975	1,432

Note: Robust standard error clustered at county level in parentheses. $*, **, ***$ indicate $p < 0.1, p < 0.05$, and, $p < 0.01$ respectively. Household characteristics includes each parent's age, age squared, educational level, the number of children, total amount of assets in a household, the number of generations in a family, and age of child. Confounding policies includes duration of confounding policies in each county. Column (1)-(3) in panel A indicate the effects by using program dummy as IV. Column (1)-(3) in panel B includes the reduced form estimates.

Appendix C

Appendix for Chapter 3

- Steps to estimate dynamic effects using AS estimators in section 3

  1. Estimate $\Delta_e(g)$ for all g and l by running the following regression:

  $$Divorce_{s,t} = \alpha_s + \gamma_t + \sum_{e \neq \infty} \sum_{g \neq -1}^{7} \Delta_e(g) \cdot 1\{E_s = e\} \cdot D^g_{s,t} + \epsilon_{s,t}.$$

  2. Calculate the average of $\Delta_e(g)$ weighted by the proportion of people in group e among people in treatment group exposed to unilateral divorce law for g periods during 1956-1988, which is

  $$\Delta(g) = \sum_{e} \frac{N_{e,g}}{N_g} \cdot \Delta_e(g)$$

  where $N_{e,g}$ and N_g are the number of treated people exposed for g periods in group e and the number of treated people exposed for g periods respectively. This procedure is proposed by Abraham and Sun (2021), which averages over the group-specific estimates associated with relative period g. ‘

- Steps to estimate dynamic effects using DD estimators in section 3

  1. Calculate the average of $\Delta_e(g)$ weighted by the proportion of people in group

e among people in treatment group exposed to unilateral divorce law for g periods during 1956-1988 by using `did_multiplegt` code in Stata, which is

$$\Delta(g) = \sum_{e} \frac{N_{e,g}}{N_g} \cdot \Delta_e(g).$$

- Steps to estimate weighted average of dynamic effects from section 3 by state population in section 4

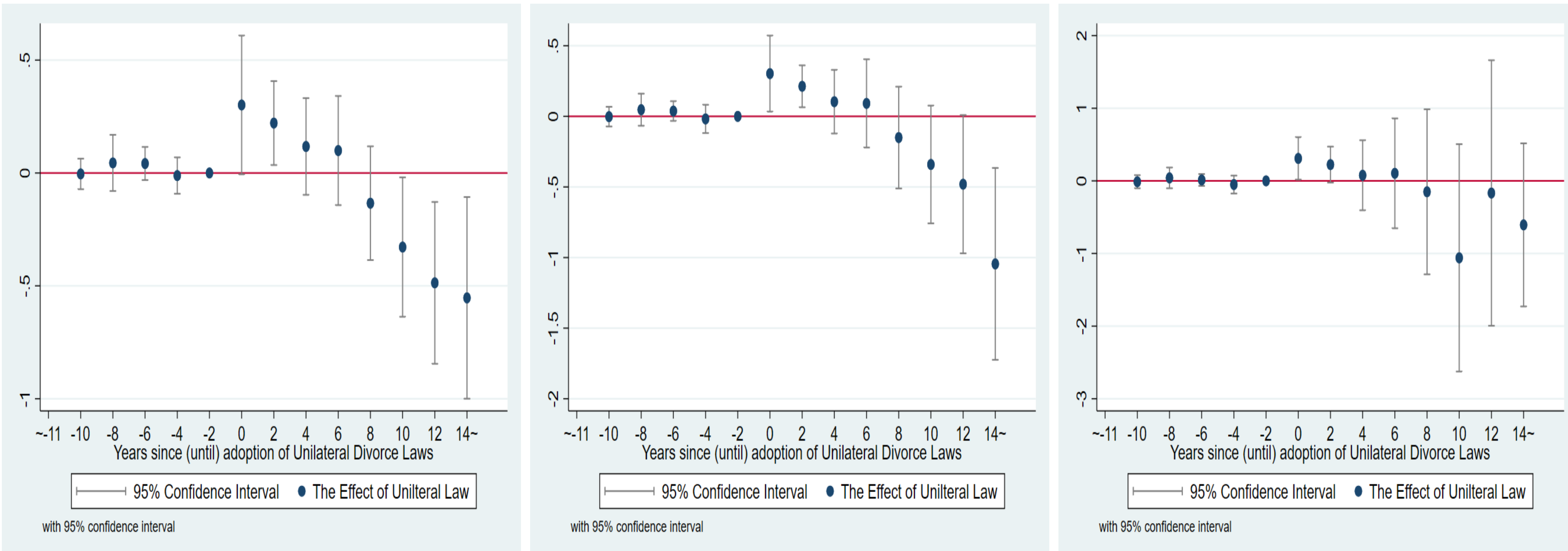
  1. Given any estimates by either AS or DD estimator, calculate the average of $\Delta(g)$ weighted by the proportion of people in all groups e exposed to the law for g periods to all people exposed to unilateral divorce law during 1956-1988, which is

     $$\sum_{g=0}^{7} \frac{N_g}{N_T} \cdot \Delta(g)$$

     where N_g and N_T are the number of treated people exposed for g periods and the number of treated people respectively.

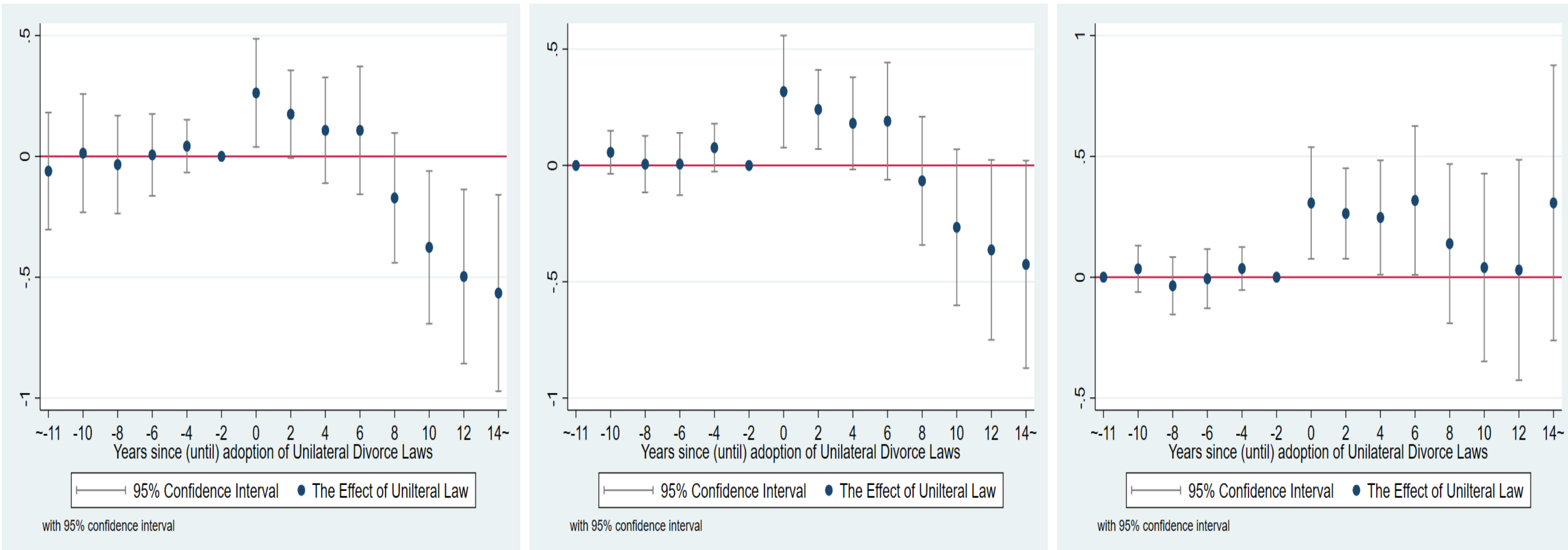
     This procedure is proposed by Callaway and SantAnna (2021), which averages all of the identified group-time average treatment effects.

Figure S1-A: Response of Divorce Rate to Unilateral Divorce Laws.



Notes: Figure S1-A shows estimates of dynamic effects using de de Chaisemartin and D'Haultfuille (2020)'s estimator. The left one indicates dynamic effects without any controls while the center one indicates the effects controlling for state-specific linear time trends. Finally, the right one shows the dynamic effects controlling for state-specific linear and quadratic time trends. Standard errors are calculated by bootstrap 500 times.

Figure S1-B: Response of Divorce Rate to Unilateral Divorce Laws.



Notes: Figure S1-B shows estimates of dynamic effects using Abraham and Sun (2021)'s estimator. The left one indicates dynamic effects without any controls while the center one indicates the effects controlling for state-specific linear time trends. Finally, the right one shows the dynamic effects controlling for state-specific linear and quadratic time trends.

Table S1: Year of The Adoption of Unilateral Divorce Law across States

Group: Year of Adoption	States
Group 1: 1969	KS
Group 2: 1970	CA, IA
Group 3: 1971	AL, CO, FL, ID, ND, NH
Group 4: 1972	KY, MI, ME
Group 5: 1973	AZ, CT, GA, HI, IN, ME, NM, NV, OR, WA
Group 6: 1974	MN, TX
Group 7: 1975	MA, MT
Group 8: 1976	RI
Group 9: 1977	WY
Group 10: 1985	SD
Never treated group:	AR, DC, DE, IL, LA, MD, MO, MS, NC, NJ NY, OH, PA, SC, TN, UT, VA, VT, WI, WV

Note: Never treated group consists of states that did not adopt unilateral divorce law during 1956-1988.

Bibliography

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2022). When Should You Adjust Standard Errors for Clustering? *The Quarterly Journal of Economics*, 138(1):1–35.

Abraham, S. and Sun, L. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Acemoglu, D., Autor, D., Dorn, D., Hanson, G. H., and Price, B. (2016). Import Competition and the Great US Employment Sag of the 2000s. *Journal of Labor Economics*, 34(S1):S141–S198.

Allen, D. W. (1992). Marriage and divorce: Comment. *The American Economic Review*, 82(3):679–685.

Andrews, I., Stock, J. H., and Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11(1):727–753.

Asquith, B., Goswami, S., Neumark, D., and Rodriguez-Lopez, A. (2019). U.S. job flows and the China shock. *Journal of International Economics*, 118:123–137.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, 103(6):2121–68.

Autor, D. H., Dorn, D., Hanson, G. H., and Song, J. (2014). Trade Adjustment: Worker-Level Evidence. *The Quarterly Journal of Economics*, 129(4):1799–1860.

Balsvik, R., Jensen, S., and Salvanes, K. G. (2015). Made in China, sold in Norway: Local labor market effects of an import shock. *Journal of Public Economics*, 127:137–144. The Nordic Model.

Barr, A. and Turner, S. E. (2013). Expanding enrollments and contracting state budgets: The effect of the great recession on higher education. *The ANNALS of the American Academy of Political and Social Science*, 650(1):168–193.

Bau, N. (2021). Can policy change culture? government pension plans and traditional kinship practices. *American Economic Review*, 111(6):1880–1917.

Becker, G. S. (1981). *A Treatise on the Family*. Number beck81-1 in NBER Books. National Bureau of Economic Research, Inc.

Benguria, F. and Ederington, J. (2021). Decomposing the Effect of Trade on the Gender Wage Gap. Working Paper.

Bertrand, M., Cortes, P., Olivetti, C., and Pan, J. (2020). Social Norms, Labour Market Opportunities, and the Marriage Gap Between Skilled and Unskilled Women. *The Review of Economic Studies*, 88(4):1936–1978.

Besedes, T., Lee, S. H., and Yang, T. (2021). Trade liberalization and gender gaps in local labor market outcomes: Dimensions of adjustment in the United States. *Journal of Economic Behavior & Organization*, 183:574–588.

Betts, J. R. and McFarland, L. L. (1995). Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments. *Journal of Human Resources*, 30(4):741–765.

Borusyak, K., Hull, P., and Jaravel, X. (2021a). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies*, 89(1):181–213.

Borusyak, K., Jaravel, X., and Spiess, J. (2021b). Revisiting Event Study Designs: Robust and Efficient Estimation. Papers 2108.12419, arXiv.org.

Brandt, L., Van Biesebroeck, J., and Zhang, Y. (2012). Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics*, 97(2):339–351.

Caldwell, J. C. (1976). Toward a restatement of demographic transition theory. *Population and Development Review*, 2(3/4):321–366.

Callaway, B. and SantAnna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230. Themed Issue: Treatment Effect 1.

Campana, J. C., Gimenez-Nadal, J. I., and Molina, J. A. (2018). Gender Norms and the Gendered Distribution of Total Work in Latin American Households. *Feminist Economics*, 24(1):35–62.

Charles, K. K., Hurst, E., and Notowidigdo, M. J. (2018). Housing Booms and Busts, Labor Market Opportunities, and College Attendance. *American Economic Review*, 108(10):2947–94.

Cheng, L., Liu, H., Zhang, Y., and Zhao, Z. (2018). The heterogeneous impact of pension income on elderly living arrangements: evidence from chinas new rural pension scheme. *Journal of Population Economics*, 31(1):155–192.

Chiappori, P.-A., Iyigun, M., and Weiss, Y. (2015). The Becker-Coase Theorem Reconsidered. *JODE - Journal of Demographic Economics*, 81(2):157–177.

Dahl, G. B. and Moretti, E. (2008). The Demand for Sons. *The Review of Economic Studies*, 75(4):1085–1120.

Dauth, W., Findeisen, S., and Suedekum, J. (2014). The Rise of The East and The Far East: German Labor Markets and Trade Integration. *Journal of the European Economic Association*, 12(6):1643–1675.

de Chaisemartin, C. and D'Haultfuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

Dobkin, C., Finkelstein, A., Kluender, R., and Notowidigdo, M. J. (2018a). The economic consequences of hospital admissions. *American Economic Review*, 108(2):308–52.

Dobkin, C., Finkelstein, A., Kluender, R., and Notowidigdo, M. J. (2018b). The economic consequences of hospital admissions. *American Economic Review*, 108(2):308–52.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813.

Eagan, K. (2016). *The American Freshman: Fifty-year Trends, 1966-2015*. Higher Education Research Institute, Graduate School of Education & Information Studies, University of California, Los Angeles.

Ebenstein, A. and Leung, S. (2010). Son preference and access to social insurance: Evidence from china's rural pension program. *Population and Development Review*, 36(1):47–70.

Fadlon, I. and Nielsen, T. H. (2019). Family health behaviors. *American Economic Review*, 109(9):3162–91.

Ferriere, A., Navarro, G., and Reyes-Heroles, R. (2021). Escaping the Losses from Trade: The Impact of Heterogeneity on Skill Acquisition. Working Paper.

Foote, A. and Grosz, M. (2020). The Effect of Local Labor Market Downturns on Postsecondary Enrollment and Program Choice. *Education Finance and Policy*, 15(4):593–622.

Friedberg, L. (1998). Did unilateral divorce raise divorce rates? evidence from panel data. *The American Economic Review*, 88(3):608–627.

Giuliano, P. (2021). Gender and culture. *Oxford Review of Economic Policy*, 36(4):944–961.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277. Themed Issue: Treatment Effect 1.

Greenland, A. and Lopresti, J. (2016). Import exposure and human capital adjustment: Evidence from the U.S. *Journal of International Economics*, 100:50–60.

Han, B. and Kang, H. (2013). History of Financial Support for Human Capital with Emphasis on Student Aid Policy (in Korean). *The Review of Business History*, 28(3):45–79.

Helm, I. (2019). National Industry Trade Shocks, Local Labour Markets, and Agglomeration Spillovers. *The Review of Economic Studies*, 87(3):1399–1431.

Heß, S. (2017). Randomization Inference with Stata: A Guide and Software. *The Stata Journal*, 17(3):630–651.

Ho, C. (2019). Childs gender, parental monetary investments and care of elderly parents in China. *Review of Economics of the Household*, 17(3):741–774.

Hsieh, C.-T. and Ossa, R. (2016). A global view of productivity growth in China. *Journal of International Economics*, 102:209–224.

Huang, W. and Zhang, C. (2021). The power of social pensions: Evidence from china's new rural pension scheme. *American Economic Journal: Applied Economics*, 13(2):179–205.

Hubbard, D. A. (2018). The Impact of Local Labor Market Shocks on College Choice: Evidence from Plant Closings in Michigan. Working Paper.

Hwang, J. (2016). Housewife, "gold miss," and equal: the evolution of educated women's role in asia and the u.s. *Journal of Population Economics*, 29(2):529–570.

Hwang, J., Lee, C., and Lee, E. (2019). Gender norms and housework time allocation among dual-earner couples. *Labour Economics*, 57:102–116.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Jaravel, X., Petkova, N., and Bell, A. (2018). Team-specific capital and innovation. *American Economic Review*, 108(4-5):1034–73.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics*, 7(1):63–88.

Kim, D., Huh, M., and Lee, D. (2007). Analysis of the Regional Economy by Defining Korean MSAs (in Korea).

Kneip, T., Bauer, G., and Teachman, J. (2009). Did unilateral divorce laws raise divorce rates in western europe? *Journal of Marriage and Family*, 71(3):592–607.

Lee, J. Y. and Solon, G. (2011). The Fragility of Estimated Effects of Unilateral Divorce Laws on Divorce Rates. *The B.E. Journal of Economic Analysis & Policy*, 11(1):1–11.

Lee, M. J. (2021). The effect of import competition on educational attainment at the post-secondary level: Evidence from NAFTA. *Economics of Education Review*, 82:102117.

Li, H., Li, L., Wu, B., and Xiong, Y. (2012). The End of Cheap Chinese Labor. *The Journal of Economic Perspectives*, 26(4):57–74.

Liu, H., Sun, Q., and Zhao, Z. (2014). Social learning and health insurance enrollment: Evidence from china's new cooperative medical scheme. *Journal of Economic Behavior and Organization*, 97:84–102.

Liu, K. (2016). Insuring against health shocks: Health insurance and household choices. *Journal of Health Economics*, 46:16–32.

Lochner, L. and Monge-Naranjo, A. (2012). Credit Constraints in Education. *Annual Review of Economics*, 4(1):225–256.

Lovenheim, M. F. (2011). The Effect of Liquid Housing Wealth on College Enrollment. *Journal of Labor Economics*, 29(4):741–771.

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

Lusardi, A. and Mitchell, O. S. (2007). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics*, 54(1):205–224. Carnegie-Rochester Conference Series on Public Policy: Economic Consequences of Demographic Change in a Global Economy April 21-22, 2006.

Naughton, B. (2007). *The Chinese Economy: Transitions and Growth*, volume 1 of *MIT Press Books*. The MIT Press.

Ok, W., Jung, S., and Oh, Y. (2007). Trade, International Division of Labor, Skill-Demand Structure and Wage Inequalities: With a Focus on the Korea-China Trade (in Korean). *Journal of Korean Economic Analysis*, 13(3):73–135.

Olea, J. L. M. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369.

Oliveira, J. (2016). The value of children: Inter-generational support, fertility, and human capital. *Journal of Development Economics*, 120:1–16.

Peters, H. E. (1986). Marriage and Divorce: Informational Constraints and Private Contracting. *American Economic Review*, 76(3):437–454.

Peters, H. E. (1992). Marriage and Divorce: Reply. *American Economic Review*, 82(3):687–693.

Poterba, J. M., Venti, S. F., and Wise, D. A. (1995). Do 401(k) contributions crowd out other personal saving? *Journal of Public Economics*, 58(1):1–32.

Shen, Z., Zheng, X., and Yang, H. (2020). The fertility effects of public pension: Evidence from the new rural pension scheme in china. *PLOS ONE*, 15(6):1–17.

Shenhav, N. (2021). Lowering Standards to Wed? Spouse Quality, Marriage, and Labor Market Responses to the Gender Wage Gap. *The Review of Economics and Statistics*, 103(2):265–279.

Solon, G., Haider, S. J., and Wooldridge, J. M. (2015). What are we weighting for? *The Journal of Human Resources*, 50(2):301–316.

Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology*, 35(1):4–28.

Steele, C. M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. *The American psychologist*, 52 6:613–29.

Stepan, M. and Lu, Q. (2016). The establishment of china's new type rural social insurance pension: A process perspective. *Journal of Current Chinese Affairs*, 45(2):113–147.

Tang, C., Zhao, L., and Zhao, Z. (2020). Does free education help combat child labor? The effect of a free compulsory education reform in rural China. *Journal of Population Economics*, 33(2):601–631.

Wang, D. (2006). China's Urban and Rural Old Age Security System: Challenges and Options. *China & World Economy*, 14(1):102–116.

Wolfers, J. (2006). Did unilateral divorce laws raise divorce rates? a reconciliation and new results. *American Economic Review*, 96(5):1802–1820.

Yang, J. and Chen, X. (2022). Grandfathers and Grandsons: Social Security Expansion and Child Health in China. IZA Discussion Papers 15239, Institute of Labor Economics (IZA).

Zafar, B. (2013). College Major Choice and the Gender Gap. *The Journal of Human Resources*, 48(3):545–595.

Zhang, Z., Luo, Y., and Robinson, D. (2019). Who are the beneficiaries of chinas new rural pension scheme? sons, daughters, or parents? *International Journal of Environmental Research and Public Health*, 16(17).

Zhao, Q., Brosig, S., Luo, R., Zhang, L., Yue, A., and Rozelle, S. (2016). The new rural social pension program in rural china: Participation and its correlates. *China Agricultural Economic Review*, 8:647–661.

Zhu, X. (2012). Understanding China's Growth: Past, Present, and Future. *Journal of Economic Perspectives*, 26(4):103–24.