Demographic Transition and Structural Transformation in China

Ningyuan Jia

LSE

October 4, 2023

Abstract

This paper examines the relationship between demographic transition and structural transformation in China from 1982 to 2000. I develop a dynamic economic geography model that incorporates dynamic fertility decision-making, multi-region and multi-sector settings, non-homothetic preferences, technology progress, and China-specific frictions. The mechanism involves changes in the total labour force along with disparities in sectoral factor intensities. After the implementation of One Child Policy (OCP), the total labour force expands in the short run due to a lower dependency ratio but contracts in the long run due to decreased fertility. Labour is channeled more towards the labor-intensive non-agricultural sector compared to the land-intensive agricultural sector due to factor intensity differences, resulting in structural transformation. Empirical analysis utilizing regional variations in OCP confirms this relationship and its underlying mechanism. Counterfactual analysis from the model reveals that in the short run, the OCP leads to a 0.91 percentage point increase in the share of non-agricultural employment. However, in the long run, it results in a 2.7 percentage point decrease in the share of non-agricultural employment.

1 Introduction

The demographic transition and structural transformation are two crucial processes that have significant implications for economic development and societal changes in many countries. Demographic transition refers to the shift from high fertility and mortality rates to low fertility and mortality rates, which opens a demographic window of opportunity for industrialization and urbanization. Structural transformation, on the other hand, involves the shift in economic activities from agriculture to industry and services, which is essential for fostering industrialization and improving living standards. Many countries have implemented family planning programs in their early development stages, aiming to seize the demographic window and foster industrialization. By the mid-1990s, such programs were active in 115 countries[1]. During this period, world total fertility rate in developing countries fell by more than half. While there is a large body of literature focusing on family-level outcomes of such programs, the linkage between demographic transition and structural transformation is lacking. Moreover, we know very little about the the trade-off between the short-run dependency ratio and long-run labour force due to changes in the age structure.

My research focuses on China as an ideal case study to explore the relationship between demographic transition and structural transformation. Two critical facts make China unique for this investigation. Firstly, China has undergone a profound economic evolution, transitioning from an agriculture-focused economy to a manufacturing-driven model. Over the decades, the share of agricultural sector employment in China exhibited a significant decline, dropping from 70% in 1980 to 50% in 2000 (as depicted in Figure 1). By 2010, this figure had further reduced to 39%. While numerous influential factors contributed to this transition, my specific focus is on one pivotal factor: the demographic transition.

Secondly, the implementation of a series of family planning programs during the 1970s and the introduction of the One Child Policy (OCP) in the early 1980s have exerted substantial impacts on China's population dynamics. Figure 2 illustrates the fertility trend, which represents the number of children per mother from 1960 to 2000. The significant decrease in fertility is closely associated with the extensive implementation of these family planning programs. Moreover, this decline in fertility rates has induced profound changes not only in the fertility rate itself but also in the age structure of the population and the dependency ratio. Figure 3 compares the population pyramid between 1982 and 2000. In 1982, the population pyramid displayed a higher proportion of individuals under 15, indicative of a youthful population. By 2000, the age structure shifted, with a greater share of individuals between 15 and 65, reflecting a transition towards a higher share of working-age population. This change has significant implications for dependency ratios, as measured by the number of young dependents (under 15) over total working age population who aged between 15 and 64 in Figure 4, which shows a consistent decline in the dependency ratio. This paper aims to investigate how changes in age structure resulting from the OCP influence

[1] For a comprehensive review of global family planning programs, see Robinson and Ross [2007]

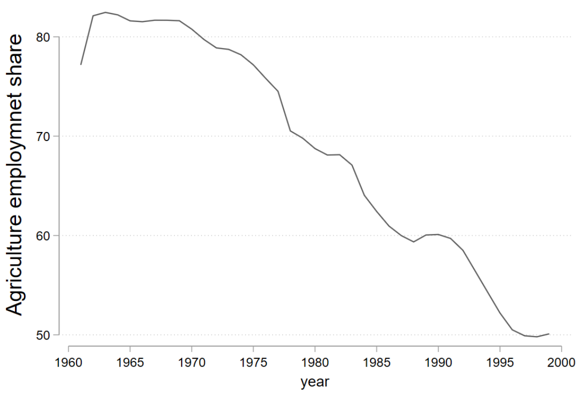

Figure 1: Industrilization in China

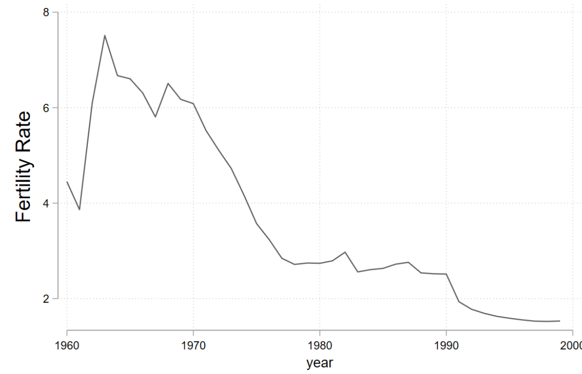

Figure 2: Fertility Trend in China

short-term and long-term structural transformation. By delving into China's developmental trajectory, My paper offers inspiration for policies on demographic and economic transitions.

In this paper, I examine how demographic transition caused by OCP has influenced the process of structural transformation in China. The paper consists of three parts. First, I develop a dynamic economic geography model that incorporates the key mechanism underlying demographic transition: differences in sectoral factor intensity coupled with shifts in the labor force size. Second, I use exogenous variation in the OCP shock across prefectures to provide suggestive evidence on the relationship between demographic transition and structural transformation and empirically evaluate the channels depicted in the model. Third, I take the model to data and perform counterfactual analysis by removing OCP. This model enables me to examine the driving forces behind structural transformation while considering the demographic transition's impact and assess aggregate welfare effects across regions. By adding China specific friction and other comment drivers in the model, it also enables me to distinguish the contributions of OCP from other traditional forces including technology change, entry barriers to non-agricultural sector, migration cost, trade cost and Hukou restrictions[2].

The dynamic economic geography model in my paper aims to analyze the effect of the OCP, understand its channel, quantify its aggregate impact, and study its interaction with other traditional forces. The model features the following key components: (1) dynamic fertility decision-making, (2) standard multi-region and multi-sector urban model component, (3) non-homothetic preferences and technology progress in different sectors, and (4) China-specific frictions: Hukou cost and non-agricultural entry barriers. To this end, I incorporate endogenous fertility in line with canonical theory in family economics. The model also features

[2]In China, "Hukou" refers to the household registration system. It is a crucial administrative system that identifies and records individuals' official residential status in the country. Every Chinese citizen is assigned a Hukou, which can be either agricultural (rural) or non-agricultural (urban) in a specific administrative unit. Bian [2002] and Chan [2010] discuss the Hukou system in detail.

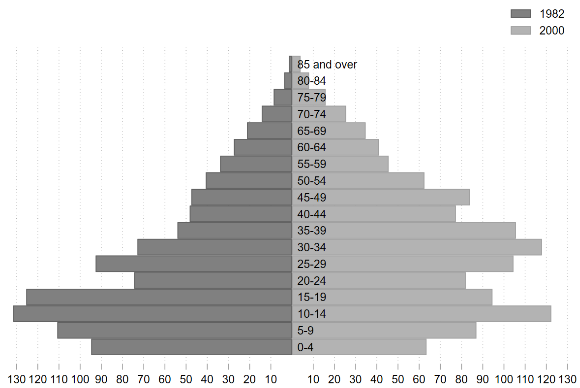

Figure 3: Population Pyramid (in Millions)

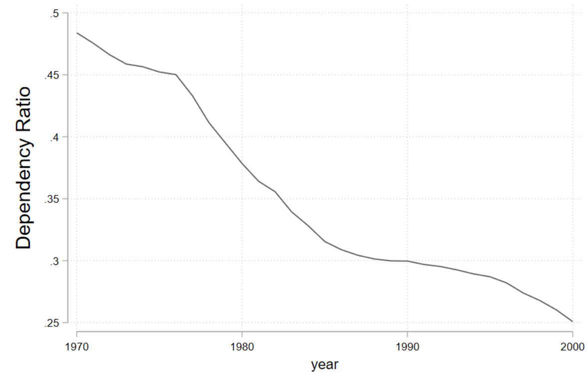

Figure 4: Dependency Ratio

overlapping generations setting, and an OCP-specific shock is added to capture China's unique policy environment. This enables me to match China's specific demographic and economic characteristics and investigate the key channels identified in the empirical findings.

In my model, individuals are assumed to live for two periods. During the first period, they are children and do not participate in any decision-making. In the second period, they enter the labour market and make work and fertility choices. When individuals decide to raise children, they not only need to allocate some time out of their working hours to take care of their children, but they also face OCP-specific punishment if they have more children than allowed by the policy. Additionally, individuals derive utility from having children, and this utility is based on the expected future utility of their offspring, as expressed in Barro and Becker [1989].. On the production side, each location is characterized by two sectors: agricultural sector and non-agricultural sector. The agricultural sector is modeled as land-intensive and operates under perfect competition, while the non-agricultural sector is modeled as labour-intensive and operates under monopolistic competition with region-specific entry costs.

The intuition of the OCP effect in my model is as follows: an increase (decrease) in the OCP's punishment for having more children would raise (lower) the cost of raising a child and in turn reduce (increase) fertility rates. Consequently, individuals would mechanically spend less (more) time on child-rearing. Since children do not enter the labour market in the first period, a change in fertility rates directly affects the size of the total labour force in the short run. When family planning programs are implemented, reduced fertility and a decreased dependency ratio contribute to an increase in the total effective labour supply in the short run. This effect is often referred to as a "demographic window of opportunity," wherein a lower dependency ratio allows for a potentially more productive and economically active labour force, which can promote economic growth and industrialization. In the second period of the model, when the newborns from the first period enter the labour market, the total labour force decreases (increases) in response to the initial increase (decrease) in the OCP's punishment. This happens because the decrease (increase) in fertility caused by the OCP's punishment leads to a smaller (larger) cohort of new entrants into the labour force in the second period. Since labour-intensive (non-agricultural sector) sector expend more when there is an increase in total labour supply, The structural transformation pattern can be attributed to the combination of differences in factor input intensity in two sectors and the reverse trend of total labour supply in the short run and long run.

The geographic setting of the model plays a crucial role in ensuring the basic intuition of the OCP effect. (i) The presence of migration frictions implies that a significant share of newborns can only be absorbed locally, limiting their mobility across regions. If there were no migration costs, individuals would be able to migrate frictionlessly to other regions, and local shocks would have a minimal impact on the labour force dynamics in each region. In such a scenario, a shock affecting one region would be equivalent to a shock affecting all regions uniformly, and there would be no discernible difference between regions with different shocks. The consideration of migration frictions in the model is essential for capturing the localized impact of the OCP on demographic changes and structural transformation. In the OCP setting, the geographic restrictions on migration imply that an increase in fertility in a certain region would primarily affect the local labour supply, and this effect may not be immediately offset by migration from other regions. As a result, regions with different OCP policy exposures may experience varying changes in fertility rates and labour force sizes, leading to differences in their economic and industrial development trajectories over time. (2) The presence of inter-regional trade is also a crucial aspect of the model, impacting the interpretation of the OCP effect. In the absence of trade with other regions, an increase in local labour supply would lead to a greater boost in output in the relatively labour-intensive sector (non-agricultural sector) compared to the land-intensive sector (agricultural sector), assuming constant prices. As a result, the relative price of agricultural products and wages would rise, attracting more labour into the agricultural sector. However, when trade is considered, local sectoral prices become influenced by costs in other regions and are less affected by local output changes. This dampens the equilibrium price response to regional shocks, thereby allowing the factor input intensity effect to play a dominant role in shaping the structural transformation pattern.

Along with traditional drivers of structural transformation such as non-homothetic preferences and technological progress, my model comprehensively captures the majority of important factors that have influenced China's industrialization since the reform and opening up in 1978. It accounts for rapid technological advancements in both sectors, lower entry barriers to the non-agricultural sector, reduced Hukou and trade costs, and OCP. Through these features, the model enables me to disentangle the specific effect of the OCP from other factors, study their interactions, and accurately assess the contribution of each factor to the overall process of structural transformation in China's economy.

After exploring the key channels and qualitative insights in the model, I use OCP as a natural experiment to provide suggestive evidences for the mechanism and predictions behind the model. There are several difficulties in identifying the effect of demographic transition using OCP. One of the main difficulties is that OCP was not randomly assigned, which means that the groups of people who were affected by the policy may differ in other ways that affect the outcomes of interest. Additionally, the implementation of the OCP at the regional level was based on local socioeconomic considerations, which means that the variation in OCP enforcement partly comes from variation in regional fundamentals. This creates challenges in exploring exogenous variation and comparing the outcomes across different groups of people and regions.

To address the endogeneity difficulties, I exploit the relaxation in OCP after 1982 in most provinces which allowed families to have a second child if the first child is a girl. Other families were still only allowed to have one child. Therefore, the affected groups are single-daughter families with only one child, and that child being a girl after the relaxation. Consequently, regions with more single-daughter families received larger policy shock. Since there is no evidence for gender-selection at the first birth in 1982, which I confirm in my empirical study, the gender of the first birth is plausibly exogenous and is not associated with local economic conditions. In the main identification, I use the share of single-daughter families in 1982 as the measure of policy exposure in each region. To avoid families intentionally abandoning girls in their early years before the relaxation to have more boys, I only consider families that gave birth to their first child in 1982 when calculating the share. The average share of single-daughter families is 49.5%, which translates to a sex ratio of 102 (number of boys for every 100 girls). This lies in the normal range of the child sex ratio.

The exogeneity of the gender of the first birth plays a crucial role in ensuring the validity of the parallel trend assumption for the DID estimation. This assumption states that, in the absence of the OCP relaxation, the trends in outcome variables should be parallel across regions with different policy exposures. In other words, regions with higher and lower shares of single-daughter families should have experienced similar trajectories in employment outcomes had they been subject to the same economic and policy conditions. To provide evidence supporting the parallel trend assumption, I conduct a comprehensive analysis of pre-treatment trends in employment outcomes in both sets of regions. I examine the employment data for agricultural and non-agricultural sectors from the years leading up to the OCP relaxation, before any differential policy effects could have taken place.

My DID estimation compares changes in agricultural and non-agricultural employment between prefectures with different shares of single-daughter families before and after the 1982 relaxation of the OCP using data from the 1982, 1990, and 2000 1% population censuses. The baseline results indicate that regions with a higher share of single-daughter families experienced a greater increase in non-agricultural employment share during 1982-2000, suggesting a faster pace of industrialization during that time. However, in the short run (1982-1990), these same regions witnessed a decrease in non-agricultural employment share or slower industrialization. These contrasting outcomes shed light on the complexities of the connection between demographic changes and structural transformation.

Why would the relaxation contribute to structural transformation, and why does the trend reverse in the short and long run? For any possible factors to account for the reverse trend in the main findings, they need to satisfy three criteria. (1) It should be affected by the share of single-daughter families. (2) It should also have different trends in the short and long run. (3) It is associated with structural changes. The main explanation in this paper is the total labour supply combined with differences in factor intensity in the agricultural and non-agricultural sectors.

By replacing the outcome variable with total labour force in the baseline regression, I find that the total labour supply shows different trends in the short run and long run due to changes in fertility and the dependency ratio following the relaxation of the OCP. The most direct effect of the relaxation is that it led to higher fertility rates in regions with high shares after 1982. It takes at least 15 years for these newborns to enter the labour market; therefore, the total labour supply only increased in these regions by the year 2000 and not in 1990. Another aspect of the increase in fertility is the rise in the dependency ratio, which indeed led to a decrease in effective labour supply in 1990. This is because more families needed to take time off from their working hours to take care of their children, reducing the available labour force in the short run.

These changes in fertility and the dependency ratio contributed to the observed trend reversal in the short run and long run, impacting the overall labour supply and influencing the process of structural transformation in the economy. Specifically, in China at 2000, the agricultural sector was more land-intensive, while the non-agricultural sector was more labour-intensive. As a result, an increase in the local labour supply would have led to a greater proportion of labour working in the non-agricultural sector compared to the agricultural sector when regions are interconnected through trade.

Apart from the total labour force channel, the structural transformation may also be affected through other factors, such as the quantity-quality trade-off and the sex ratio. Regions with higher shares of single-daughter families might have experienced changes in parental investment and educational opportunities. This improvement in human capital could have contributed to the long-term process of structural transformation, as a better-educated and skilled workforce is essential for the growth of industries and the economy. Additionally, the sex ratio imbalance resulting from the OCP might have led to a surplus of males in certain regions. This could also affect structural transformation if males and females have different propensities toward different sectors. In this paper, I empirically assess these channels and demonstrate that they are not the main drivers behind the findings in China's context.

After providing empirical evidence on the key mechanism, I take the model to data to quantify the aggregate effect of OCP on structural transformation. The quantitative results from the model show that OCP had divergent effects on China's industrialization. In the short run during 1982-1990 , industrialization at the national level experienced a 0.91% faster growth rate. However, the long-run consequences of reduced labour supply during 1982-2000 due to the OCP led to an overall 2.73% slower industrialization. These effects were more pronounced in regions with high relative non-agricultural Total Factor Productivity (TFP), low Hukou costs, low trade costs, and high migration costs. Furthermore, Hukou cost reduction and TFP increase are identified as major drivers of China's overall structural transformation. Though real wage increased in the long run, welfare decreased largely due to the lower utility from children arising from high OCP fine.

Related literature. My paper mainly contributes to three strands of literature. The first one is the family economic literature in understanding the causes and consequences of fertility (Doepke et al., 2022; Greenwood et al., 2021; Greenwood et al., 2017; Doepke and Tertilt, 2016;Jones et al., 2008;Hotz et al., 1997). This literature identifies two major drivers of fertility to account for the observed negative income-fertility relationship in most countries including China. The first one is the quantity quality trade off (Delventhal et al., 2021; Croix and Licandro, 2013; Doepke and Tertilt, 2009; Manuelli and Seshadri, 2009; Moav, 2005; Doepke, 2005; Becker, 1960). However, in the context of China, the evidence on the quantity quality trade off is mixed. Qian [2009] finds that having an additional child increased the likelihood of school enrollment of the first child by about 16 percentage points, suggesting a complementary relationship between quantity and quality in education. Rosenzweig and Zhang [2009] found a positive but a modest effect of the quantity on quality. My empirical result adds to the literature and shows on the regional level, this is no evidence for the quantity-quality trade-off. Another driver for fertility is the opportunity cost of women's time (Coskun and Dalgic, 2022; Currie and Schwandt, 2014; Hotz et al., 1997; Schultz, 1997 Galor and Weil, 1993). One effect of this channel is the negative relationship between fertility and female labour force participation (Aaronson et al., 2021). In this paper, I show that the time cost of raising children is key to understanding the changes in fertility in response to the family planning program shocks and main driver behind the reverse trend in total labour supply.

The second strand of literature is the macroeconomic literature in understanding drivers of structural transformation. This literature explores contribution of various factors including technological progress, frictions, non-homothetic tastes, capital intensities, substitutability, trade on structural transformation (Erten and Leight, 2021; Fajgelbaum and Redding, 2022;Ngai and Petrongolo, 2017; Święcki, 2017; Herrendorf et al., 2014; Bar and Leukhina, 2010; Buera and Kaboski, 2012; Brandt et al., 2008; Ngai and Pissarides, 2007). However, the literature has not focused on the role of population growth. Leukhina and Turnovsky [2016a] and Leukhina and Turnovsky [2016b] provide a general equilibrium model analysing the effect of population growth on structural change and show that the population effects depend on the preference-side and production-side characteristics of the economy, and trade. Barham et al. [2023] find that family planning program in Bangladesh slows down structural transformation. Models in this literature miss the spatial dimension and focus on the aggregate economy. So they can not explain the regional variations in family planning programs. This is important because regional variations allow me to hold national level policies constant when assessing the impact of family planning programs. This could be valuable in many countries where several national policies are adopted together to promote growth. Ngai et al. [2019] highlight spatial dimension using China's Hukou system, studying its effect on industrialization. To this end, the detailed mechanism underlining the population effect on a regional level is still unclear and the short-term and long-term trade-offs remain unexplored.

The third strand is the urban and trade literature on the quantitative spatial model (Fan et al., 2021; Eckert and Peters, 2022a; Desmet et al., 2018; Caliendo and Parro, 2021; Caliendo et al., 2019; Redding and Rossi-Hansberg, 2017; Allen and Arkolakis, 2014). Within this fast growing literature, some of them focus on the spatial dimension of structural transformation. For example, Eckert and Peters [2022a] propose a spatial model of structural transformation with the key driver being catch-up growth. As in most spatial models, population growth is assumed to be exogenous. My model incorporates endogenous fertility decision making that allows me to study the effect of fertility related shocks on structural change.

This paper proceeds as follows: Section 2 describes the institutional background. Section 3 presents the structural model. Section 4 contains the reduced-form analysis and robustness tests as well as mechanisms. Section 5 presents the calibration and counterfactual analysis. Section 6 concludes the paper.

2 Background

2.1 One Child Policy

The One Child Policy in China was implemented in 1979 as a major population control measure to address the country's rapid population growth. Prior to the 1970s, China's population had been expanding rapidly, leading to concerns about resource scarcity and economic strain. The idea of implementing a population control policy had been considered earlier, but it was only in 1979 that the OCP was formally introduced.

Under the OCP, almost all families in China were restricted to having only one child, and second or higher-parity births were penalized. The policy was rigorously enforced, and measures such as forced abortions, fines, and other penalties were used to ensure compliance. In the initial years of the policy, there were significant efforts to control population growth, and the fertility rate in China declined as a result.

A crucial turning point occurred in 1982 when the Chinese government included the One Child Policy in the national constitution. By doing so, the OCP became a compulsory and legally binding policy for all Chinese citizens. Its inclusion in the constitution strengthened its authority and provided a clear legal basis for its enforcement, solidifying its position as a long-term population management measure.

However, as time passed, the OCP faced growing opposition and criticism from the public. Many Chinese citizens expressed dissatisfaction with the strict one-child policy, especially in rural areas, where larger family sizes were often considered necessary for agricultural labour and caring for elderly family members. Additionally, the social preference for male children to carry on the family lineage contributed to the discontent with the policy.

In response to public concerns, the Chinese government began to gradually relax the One Child Policy (OCP) in the 1980s. Local governments took the initiative to issue permits for a second child as early as 1982. However, the most significant relaxation occurred with the issuance of "Document 7" by the Central Party Committee on April 13, 1984. This document served two main purposes: first, to curb female infanticide, forced abortion, and forced sterilization; and second, to devolve responsibility from the central government to the local and provincial governments, allowing for regional variation in family planning policies.

The key relaxation introduced after "Document 7" was the "1-son-2-child" rule, which allowed rural couples to have a second child if their first child was a girl. Initially, only a small percentage of rural families were granted permits for second children in 1982. However, with the implementation of "Document 7," the number of permits increased significantly. By 1986, approximately 50% of rural families were allowed to have a second child if the first child was a girl. Following this relaxation, the fertility policy remained relatively stable and was rigorously implemented across the country until recent years.

After the relaxation of the One Child Policy (OCP), the implementation of the "1-son-2-child" rule had a considerable scope across China. As documented by Huang et al. [2019], among the 31 provinces in mainland China, 25 provinces offered this exemption to at least some rural population in their respective regions. 19 provinces extended this exemption to cover all rural residents, accounting for 71.3% of the total GDP and 74.9% of the total population in mainland China.

2.2 Hukou System

The Hukou system in China is a household registration system that has played a significant role in shaping the country's socio-economic landscape. It was introduced in the 1950s as a means of controlling population movement and resource allocation during China's planned economy era (Chan, 2010). Under the Hukou system, every individual is assigned a Hukou status, either rural or urban, based on their birthplace and family background. This Hukou status is linked to various entitlements and benefits, including access to education, healthcare, and social services. (Chan, 2010)

This lead to significant barriers and costs for rural Hukou holders to work in different regions and sectors. (i) Barriers to Working in Different Sectors: the Hukou system creates obstacles for rural Hukou holders to access non-agricultural sectors. Rural migrants face occupation segregation with urban Hukou holders. They usually face hiring discrimination, wage discrimination, and have a much higher probability of facing wage arrears in non-agricultural sectors (Song and Smith, 2019). (ii) Barriers to Moving Across Regions: the Hukou system restricts geographic mobility. Relocating to different regions involves complex administrative procedures, approvals, and documentation, making it difficult for individuals, especially rural Hukou holders, to seek better economic opportunities or job prospects (Solinger, 1999).Rural Hukou holders receive land for cultivation but lose land rents when migrating due to land market frictions. Access to social services, like education and healthcare, also depends on local Hukou status (Ngai et al., 2019).

3 The Model

In this section, I develop a dynamic geographic model to study the effect of the OCP and the underlying mechanisms: the changes in total labour force combined with sector factor intensity differences. While my model retains the traditional driving forces of structural transformation, such as technological change and non-homothetic preferences, it goes beyond existing spatial structural change models, for example Eckert and Peters [2022b], by integrating the endogenous dynamic fertility decision-making framework from family economics. This feature allows for a nuanced examination of how fertility changes interact with labour force dynamics and economic structure.

The basic environment is as follows. Time is discrete. The economy consists of N regions indexed by i, n, Each region stands for one prefecture in China. There are Two sectors in each region indexed by j, k. agricultural sector is perfect competitive indexed by A with a continuum of firms in each region. non-agricultural sector is monopolistically competitive indexed by M and characterised by an entry cost. There are competitive labour markets in each region sector subject to migration costs, meaning that labour can move across regions and sectors subject to a cost. Individual has sector specific Hukou indexed by $h \in \{A, M\}$ where A means rural Hukou (agricultural Hukou) and M means urban Hukou (non-agricultural Hukou). The Hukou sector affects individual's cost of shifting between sectors as well as migration cost across regions. Regions are different in term of sectoral technology as well as fundamental amenity.

3.1 Timing

I consider an overlapping generation (OLG) framework . In this framework, the population is composed of individuals belonging to different generations, each living through two distinct life periods: childhood and adulthood. These individuals are characterized by their place of birth (region i) and Hukou status (h), which are predetermined attributes.

At the beginning of time t, individuals of generation t are born in region i with Hukou h and spend their childhood in that region. The Hukou status is inherited from their parents and remains fixed throughout their life [3]. During their childhood, individuals do not make any decisions.

[3] This is largely true in reality but can also change in some cases. For example, there are 2 in 1000 permits in most provinces during 1990s who are allowed to change from rural Hukou to urban Hukou.

At the end of each period t, individuals observe an idiosyncratic preference shock specific to the location they are currently residing in. This shock influence their future decisions regarding migration and sectoral employment. At this point, individuals make migration decisions, selecting a new region n and a sector j based on the highest expected utility, considering preference shocks, migration costs, and Hukou restrictions as well as the expected utility they get from future fertility decision making in each region.

When the next period $(t+1)$ begins, a new generation $(t+1)$ is born in the region n where their parents migrated to. The newborn generation starts its childhood in this new location. During period $t+1$, individuals belonging to the previous generation (t) make fertility choices based on the expected utility associated with each child and the costs of child-rearing. These fertility decisions directly affect the size and composition of the next generation and, in turn, have implications for the future labour force and population distribution across regions.

At the end of period $t+1$, individuals of generation t pass away, and individuals of generation $t+1$ migrate to region n' and work in sector j'. This process of migration, sectoral employment, and fertility choices continues over time, generating dynamics in labour supply and economic outcomes in the model. I summarises the timeline below:

- At time t, generation t born in region i with Hukou h and enjoy their childhood in region i
- At the end of t, generation t migrate to region n and work in sector j with highest expected utility accounting for preference shock,migration cost and Hukou restrictions.
- At the beginning of $t+1$, they make fertility choice based on the expected utility of each child and child raising cost.
  - generation $t+1$ are born and start their childhood in region n
- At the end of $t+1$, generation t die.
  - generation $t+1$ migrate to region n' and sector j'

3.2 Preferences

The preference of each individual is determined by two components: consumption and offspring. (1) Individuals derive utility from consuming goods produced in the agricultural and non-agricultural sectors. This part of preference structure is assumed to follow a non-homothetic preference of the "PIGL" (Price-Independent Generalized Linear) class. While it do not have an explicit utility representation, the indirect utility function form make it convenient to investigate the trade-offs between consumption of goods and fertility decisions when there is a shock in the cost of raising children. (2) Individuals derive utility from the expected utility that their children will have in the next generation. This is determined by future migration decision, work decision and fertility decision of the next generation.

Raising children comes with significant costs for parents. It requires time and effort required to care for and nurture children during their childhood. This involves taking time out of the labour force to provide parental care, which can lead to a reduction in the labour supply and total income. The opportunity cost of the time spent on child-rearing is an important consideration for parents when making fertility decisions and a key driver of the labour force dynamics in the short run. Additionally, OCP also impose penalties on parents who choose to have more than the permitted number of children. One common form is a financial penalty, where parents are required to pay a fine or a share of their income as a punishment for exceeding the allowed family size.

Given these component, individual faces a trade-off between spending money on consumption and spending time on raising children. The structure of the preference is as follows:

$$U_{i,h} = (1 - k_i)(\frac{1}{\eta}(\frac{y_{i,h}}{p_{i,A}^{\phi} p_{i,M}^{1-\phi}})^{\eta} - \nu \ln(\frac{p_{i,A}}{p_{i,M}})) + k_i n_{i,h}^{\eta_f} \Pi'_{i,h}$$

where $U_{i,h}$ represents the utility of an individual living in region i with Hukou h. $\frac{1}{\eta}(\frac{y_{i,h}}{p_{i,A}^{\phi} p_{i,M}^{1-\phi}})^{\eta} - \nu \ln(\frac{p_{i,A}}{p_{i,M}}))$ is the PIGL consumption component and $n_{i,h}^{\eta_f} \Pi'_{i,h}$ is utility from offspring. k_i is a regional specific parameter that determines the weight of the offspring component in the utility function. It captures the regional-specific variation in individuals' preferences for offspring and allows me to account for the fact that different regions may have different cultural, social, and economic factors that influence individuals' fertility decisions other than the effect of OCP. $\eta \in (0, 1)$ and $\nu > 0$ are structural parameters of the PIGL function which determines the shape of Engel curve. $\phi \in (0, 1)$ is the asymptotic expenditure share on agricultural goods. $\eta_f < 1$ is a parameter that determines the individual's preference for having more children. $p_{i,A}$ and $p_{i,M}$ represent the prices of agricultural and non-agricultural goods, respectively, in region i. $\Pi'_{i,h}$ represents the expected utility that the individual's children can obtain in the future.

Given the preference, individuals in each location also have the following budget constraint to make fertility and consumption choice simultaneously

$$y_{i,h} + \chi_{i,h}(n_{i,h} - 1)I_{i,h} = (1 - qn_{i,h})I_{i,h}$$

where q represents the cost of raising a child which takes the form of a share of total working time for each child. $\chi_{i,h}$ is a penalty parameter that reflects the cost or penalty imposed on parents who choose to have more children than the permitted number allowed by the OCP. This penalty is proportional to the number of excess children beyond the permitted number [4]. $I_{i,h}$ denotes the wage of the individual in

[4] In the most of the case in China, OCP fine takes form of share of total income for each child exceed the permit. However, there are also other non-monetary punishment such as losing jobs. I aim to calibrate the model to back out the effective monetary equivalent punishment for the actually monetary and non-monetary fines

region i with Hukou status h.

With this structure at hand, Optimal fertility is given by the following equation:

$$n_{i,h}^{\eta_f - 1}(1 - qn_{i,h} - \chi_{i,h}(n_{i,h} - 1))^{1-\eta_c} = [\frac{(1-k)(q + \chi_{i,h})(\frac{I_{i,h}}{p_{i,A}^{\phi} p_{i,M}^{1-\phi}})^{\eta}}{\eta_f k \beta \Pi'_{i,h}}]$$

when $\eta = 1$, the PIGL preference reduces to a Cobb Douglas preference, then it is convinent to backout the optimal number of offspring as: $n_{i,h} = [\frac{(1-k)(q+\chi_{i,h})(\frac{I_{i,h}}{p_{i,A}^{\phi} p_{i,M}^{1-\phi}})^{\eta}}{\eta_f k \beta \Pi'_{i,h}}]^{1/(\eta_f - 1)}$ An increase in the OCP fine reduces optimal number of Children.

3.3 Income and Migration

Individuals make decisions about their location and sector choices based on expected utility, migration costs idiosyncratic preference shock. The migration process consists of two steps: location choice and sector choice upon arrival.

In the first step, individuals belonging to generation t decide whether to migrate from their current region r to another region n that offers the highest utility $\Pi_{n,h}$. This is determined by (i) The utility in region n, which is determined by the expected utility $E(U_{n,h})$, which captures the consumption and offspring utilities an individual with Hukou h can derive from living in region n (ii) amenity level A_n in region n, reflecting the attractiveness of the location, (iii) idiosyncratic preference shocks ϵ_n^i which follow an i.i.d. Fr'echet distribution $F(\epsilon_{n,i}) = e^{-(\epsilon_n^i)^{-\epsilon}}$, (iv) the migration cost from region r to region n for Hukou h denoted by $\tau_{in,h}$:

$$\Pi_{n,h} = A_n E(U_{n,h}) \frac{\epsilon_n^i}{\tau_{rn,h}}$$

Given this structure and the distribution of idiosyncratic preference shocks, the share of migrants from region r to region n for Hukou h is represented by

$$m_{rn,h} = \frac{(\bar{u}_{n,h}/\tau_{rn,h})^{\epsilon}}{\Sigma_m (\bar{u}_{n,h}/\tau_{rm,h})^{\epsilon}}$$

where $\bar{u} \equiv A_n E(U_{n,h})$ and ϵ is the parameter of Fr'echet distribution that governs the elasticity of migration across regions.

In the second step, individuals who migrate to region n make sector choices upon arrival. They select the sector j that offers the highest income, taking into account the match productivity ϕ_n^j drawn from a sector-specific Fr'echet distribution $F(\phi_i^j) = e^{-T_j \phi_i^{j-\zeta}}$ and the wage per efficient labour w_n^j in region n sector j.

Additionally, individuals with Hukou h face costs $\delta_{j,h}^n$ for working in sector j due to Hukou restrictions. These costs represent any constraints imposed on individuals with Hukou h working in specific sectors, particularly for Rural Hukou working in non-agricultural sectors.

Sectoral choice is made based on the highers income after taking the Hukou specific cost in to account: they choice sector that maximize $s = \arg\max_j\{\frac{\phi_n^j w_n^j}{\delta_{j,h}^n}\}$ and the income they get is $I_{n,h} = \phi_n^s w_n^s$.

Given this maximization structure and the distribution of idiosyncratic preference shocks, the share of individual with Hukou h working in sector j is given by upon arriving at region n:

$$s_{n,h}^j \equiv \frac{L_{n,h}^j}{L_{n,h}} = \frac{T_j(w_n^j/\delta_{j,h})^\zeta}{\Sigma_j T_j(w_n^j/\delta_{j,h})^\zeta}$$

where $L_{n,h}^j$ is the number of people in region n with Hukou h who choose to work in sector j and $L_{n,h}$ is the total number of people in region n with Hukou h. T_j is the structural parameter that determines the average match probability in sector j, ζ is the parameter of Fr'echet distribution that governs the elasticity of migration across sectors.

With migration decision $m_{rn,h}$ at hand, given the equilibrium labour supply $L_{r,h,t-1}$ at time $t-1$ in each region r for each Hukou h and the fertility decision making $n_{i,h,t-1}$ at time $t-1$, the equilibrium labour supply (generation t) at time t in each region n is given by

$$L_{n,h,t} = \Sigma_r m_{rn,h,t} L_{r,h,t-1} n_{i,h,t-1}$$

Where $L_{r,h,t-1} n_{i,h,t-1}$ represents the total number of generation $t-1$ who are still in their childhood stage and living with their parent in the same location. Entering into time t after making the migration decision $m_{rn,h,t}$, generate $t-1$ become labour force in each region where they migrate in. Summing that up over destination regions r gives the total labour supply in region r at time t for each Hukou h.

The unique aspect of this dynamic spatial model is that the fertility decisions $n_{i,h,t-1}$ made by individuals of generation $t-1$ at time $t-1$ are influenced by future utilities. This introduces a forward-looking element to the fertility decision-making process, making it more dynamic compared to other spatial models.

3.4 Production

In the non-agricultural sector, each firm produces a unique variety of non-agricultural goods denoted as ω_n^M. The production technology follows a Cobb Douglas production function with labour (L_n^M) and land (H_n^M) as inputs:

$$q_n^M(\omega_n^M) = A_n^M (L_n^M)^{\gamma_L^M} (H_n^M)^{\gamma_H^M}$$

where, A_n^M denotes the total factor productivity in the non-agricultural sector in region n and γ_L^M, γ_H^M are the share of labour and land used in production respectively, I assume a constant returns to scale technology such that they add up to 1.

Final non-agricultural goods are produced using varieties from all locations. The aggregate production of non-agricultural goods in region n is denoted as Q^{nM} and is given by the following expression:

$$Q^{nM} = (\Sigma_i \int (q_{in}^M(\omega))^{1-1/\eta^M} d\omega)^{\eta^M/(\eta^M-1)}$$

where the integral sums up the quantities of all varieties produced in different regions, and η^M represents the elasticity of substitution between different varieties. The final goods produced in the non-agricultural sector are sold only in region n.

Each variety producer in the non-agricultural sector faces monopolistic competition, which means they are able set its price. Given the demand from the final non-agricultural goods producer, the price for each variety ω_n^M produced in region i and sold in region n is given a constant markup over the unit cost c_i^M:$p_{in}^M(\omega) = \frac{\eta^M \kappa_{in}^M c_i^M(\omega)}{\eta^M - 1}$ where $\kappa_{in}^M \geq 1$ is iceberg transportation cost between region i and n.

Firms need to pay a region-specific fixed entry cost f_e^n. Free entry condition implies that profits in the non-agricultural sector are equal to the entry cost, which determines the number of firms in each region: $N_n^M(f_e^n + \gamma_L^M(\eta^M - 1)f_e^n) = L_n^M$. Given the expenditure share from consumer side, the non-agricultural goods market clearing condition gives the wage as follows:

$$w_n^M = \frac{1}{\eta^M f_e}(\frac{\eta^M}{\eta^M - 1})^{1-\eta^M} c_n^M(\omega)^{1-\eta^M} \Sigma_n[(\frac{\kappa_{in}^M}{P_n^M})^{1-\eta^M} E_n^M]$$

where E_n^M is the total expenditure in region n on sector M.

The agricultural sector follows Eaton-Kortum structure. Each location produces a continuum of agricultural varieties $\omega \in [0, 1]$ using Cobb-Douglas technology :

$$q_n^A(\omega) = z^{nA}(\omega) A_n^A (L_n^A)^{\gamma_L^A} (H_n^A)^{\gamma_H^A}$$

where, L_n^A represents the labour input in agriculture, H_n^A denotes the land input in agriculture, and $z^{nA}(\omega)$ is a productivity factor specific to each variety ω in region n. The productivity factor $z^{nA}(\omega)$ is drawn from an i.i.d. Fr'echet distribution with a cumulative distribution function given by $F(z^{nA}) = e^{-(z^{nA})^{-\theta^A}}$. The parameters γ_L^A and γ_H^A represent the shares of labour and land, respectively, in the production of agricultural goods, while A_n^A denotes the average total factor productivity in the agricultural sector in region n.

The final agricultural goods are produced by combining all varieties $\omega \in [0, 1]$ available in the market. The quantity of final goods produced in region n, denoted by Q^{nA}, is determined by aggregating the production of each variety using a CES aggregator:

$$Q^{nA} = (\int (\tilde{q}_n^A(\omega))^{1-1/\eta^A} d\omega)^{\eta^A/(\eta^A - 1)}$$

The CES aggregator combines the individual varieties $\tilde{q}_n^A(\omega)$ from different regions. The final agricultural goods are sold exclusively in region n.

The final agricultural goods producers source each variety ω from regions with the lowest price. Given the Fr'echet distribution of $z^{nA}(\omega)$, share of total expenditure in region n on agricultural varieties from region i is given by

$$\pi^{ni,A} = \frac{(\kappa_{in}^A c^{iA})^{-\theta^A} (A_i^A)^{\theta^A}}{\Sigma_{m=1}^N (\kappa_{im}^A c^{mA})^{-\theta^a} (A_m^A)^{\theta^A}}$$

where κ_{in}^A represents the iceberg transportation cost from region i to region n for agricultural varieties, and c^{iA} is the unit cost of agricultural varieties in region i. The price of final agricultural goods in region n, denoted by P^{nA}, is determined by aggregating the prices of individual varieties using a CES aggregator $P^{nA} = \Gamma^{nA}(\Sigma_{i=1}^{N}(k^{ni,A}c^{iA})^{-\theta^A}(A^{iA})^{\theta^A})^{-1/\theta^A}$

Given the expenditure share from consumer side, the agricultural goods market clearing condition gives the wage as follows:

$$w_n^A = \frac{\gamma_L^A}{L_n^A}(A^{nA})^{\theta^A}(c^{nA})^{-\theta^A}\Sigma_i(\frac{\kappa_{in}^A}{P_n^A})^{-\theta^A}E_i^A$$

4 Reduced Form Analysis

4.1 Data

The primary dataset employed in this empirical study consists of the 1% sample China population census data for the years 1982, 1990, and 2000, which has been sourced from IPUMS. With its extensive coverage, encompassing approximately 10-11 million observations for each year, this dataset provides a rich source of valuable information on various demographic characteristics of individuals in China. At the individual level, the census data includes a wide array of essential demographic attributes, such as age, gender, region of residence, region of registration, Hukou status, education level, and employment status.

Crucially, the dataset also contains valuable household-level information, allowing for the identification and linkage of children within their respective families. The relationship to the household head variable serves as a key instrument for discerning family structures within each household. This variable categorizes parents as "household head" and "spouse of the household head," and labels children as "child," . Using this variable combined with household ID, I match all children with their parents at the household level and calculate the share of single-daughter families.

While some cases may exist where children have left home and are not accounted for in the sample, it is essential to note that such occurrences predominantly pertain to grown-up children who have independently departed from their original families. However, for the purpose of my analysis, this aspect will not introduce bias into the share calculation. This is because my analysis exclusively focuses on newborns in 1982, and none of these newborns would have left their parents' households at the time of calculation.

For the outcome variables related to sectoral employment, the census provides sector information that allows for a classification of individuals into primary and secondary sectors. This also enables the exploration of heterogeneity effects within the non-agricultural sector.

The mechanism analysis and model calibration utilize additional datasets, including panel data on 285 China prefectures from the China City Statistics Yearbook (1980 onwards), the 2002 National Input-Output

(IO) table from the National Bureau of Statistics of China, province-level bilateral railway trade flow data (1985 onwards), and child-caring data from the China Health and Nutrition Survey (1991 onward).

4.2 Empirical Strategy

To investigate the effect of the OCP relaxation on structural transformation, I define the policy exposure using the share of single-daughter families as follows:

$$S_r = \frac{\sum_i B^g_{ir,1982} \cdot I[C_{ir,1982} = 1]}{\sum_i B^g_{ir,1982} \cdot I[C_{ir,1982} = 1] + \sum_i B^b_{ir,1982} \cdot I[C_{ir,1982} = 1]}$$

where $B^g_{ir,1982}$ takes the value of 1 if the household i in region r gave birth to a child in 1982 [5] and that child is a girl, and 0 otherwise. Similarly, $B^b_{ir,1982}$ takes the value of 1 if that child is a boy and 0 otherwise. $I[\cdot]$ is an indicator function that takes the value of 1 if the argument is true and 0 otherwise. $C_{ir,1982}$ is the total number of children for that household in the year 1982. The argument $I[C_{ir,1982} = 1]$ indicates households with only one child in the year 1982. The numerator represents the total number of families with only one child, and that child is a girl. The denominator represents the total number of single-child families.

This metric, S_r, allows me to quantify the prevalence of single-daughter families in each region, serving as a proxy for the policy exposure to the OCP relaxation. A higher value of S_r suggests a greater prevalence of single-daughter families and indicates a larger impact of the OCP relaxation in that region.

An important assumption for this measure being exogenous is that there is no gender selection at the first birth. This is likely to be true in practice as prenatal sex determination technologies, such as ultrasound, were not widely available in China at that time. However, the measurement is limited by the fact that only surviving children are observed in the census data. This means that the share of single-daughter families is calculated based on the gender of the child surviving at the time of the survey, which might not fully represent the true prevalence of single-daughter families if certain families engaged in gender selection or abandonment practices before the survey.

The possibility of families actively abandoning children, particularly girls, to avoid penalties under the OCP could introduce bias when calculating the share of single-daughter families based on surviving children. This may lead to an underestimation of the actual prevalence of single-daughter families. More importantly, the act of sex selection based on surviving children may introduce endogeneity concerns in the analysis, arising from the correlation between the calculated share of single-daughter families and the level of OCP enforcement. Regions with stricter OCP enforcement may be more likely to engage in

[5] As mentioned in the background, the relaxation policy started at 1984 after the document 7. I am using 1982 as the initial year because of the data availability issue. 1982 is the closest year from 1984 where I can get individual level data to calculate this share.

active abandonment of girls, resulting in a lower observed share of single-daughter families. To aviod these potential biases and endogeneity issues, I focus solely on the 1982 survey data to calculate the share of single-daughter families based on newborns in that year. In other words, I consider families that welcomed their first child in 1982 and determine the share of these families with a girl as their first child. I then compare the share of single-daughter families in 1982 using newborns with the share in the year 2000 using children aged 18. If there were no instances of active abandonment, we would expect the share to be similar for both years. By comparing the share of single-daughter families in 1982 using newborns (49.5%) with the share in the year 2000 using children aged 18 (45.4%), my analysis reveals a notable difference between the two periods. This discrepancy suggests a potential issue of active child abandonment, particularly of girls, during the implementation of the OCP.

Figure 5 presents the trend in total fertility rate (TFR) before and after the 1982 OCP relaxation. The black line represents the average TFR for regions with a share of single-daughter families above the mean, while the blue line represents regions with a share below the mean. Prior to 1982, both regions exhibited similar TFR trends. However, after 1982, regions with a higher share of single-daughter families consistently demonstrated higher TFR than those with a lower share for 15 years, indicating that the former experienced a more substantial policy shock due to the OCP relaxation. This divergence in TFR trends provides suggestive support of using this share as a measure of policy exposure.

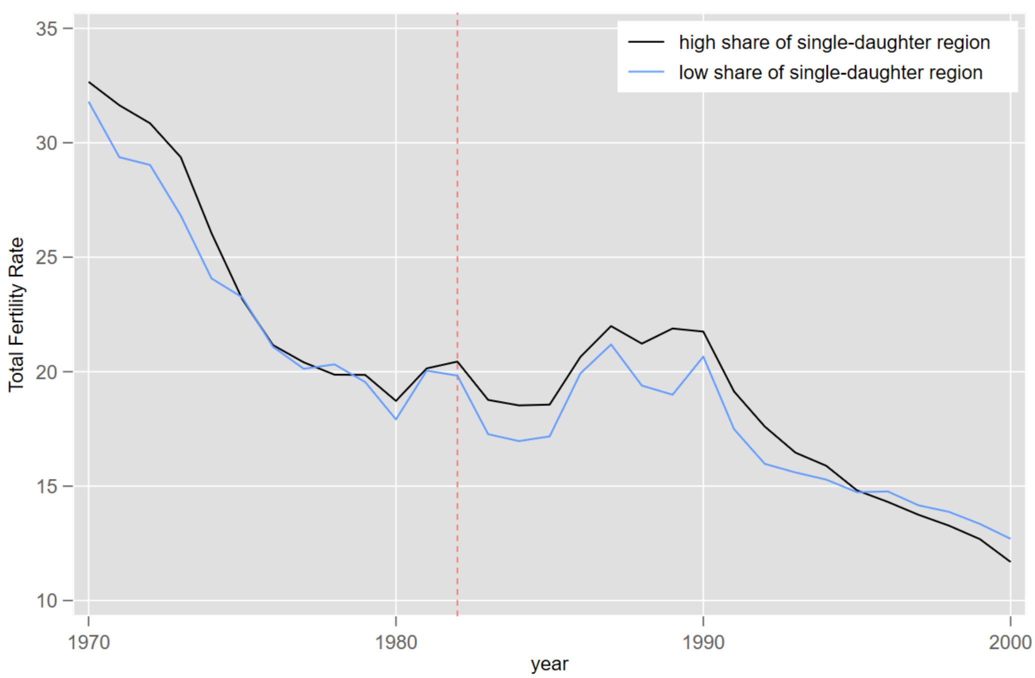

Figure 5: Total Fertility Rate Trend in Regions with Different Policy Exposure

After defining the policy exposure, the main DID estimation equation in the paper is as follows:

$$Y_{rt} = \beta \cdot post_t \cdot S_r + \delta_r + \gamma_t + \epsilon_{rt} \tag{1}$$

where Y_{rt} is the outcome variable, such as non-agricultural or agricultural employment in region r at time t. $post_t$ is a dummy variable taking the value of 1 for years after 1982 and 0 otherwise. S_r represents the share of single-daughter families in 1982. β is the key parameter of interest, capturing the effect of the One Child Policy relaxation on the outcome variable. δ_r represents region fixed effects, and γ_t denotes year fixed effects. The term ϵ_{rt} is the error term accounting for unobserved factors and random variations.

4.3 Identification Assumptions

The key identification assumption for the DID estimation is the parallel-trend assumption: in the absence of the OCP relaxation, regions with different policy exposures (i.e., varying shares of single-daughter families) would have experienced similar trends in employment outcomes under the same economic and policy conditions. This assumption is vital because it allows me to attribute any divergent changes in employment outcomes post-OCP relaxation to the policy change itself, rather than pre-existing disparities among regions.

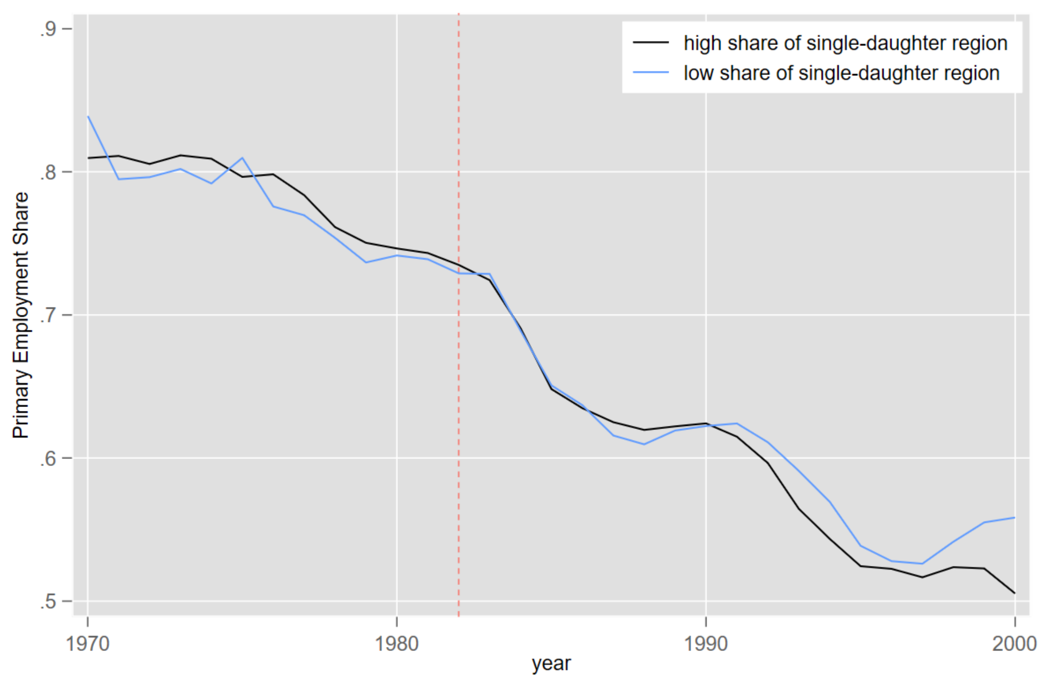

Figure 6: Primary Employment Share Trend in Regions with Different Policy Exposure

Since prefecture-level employment data is not available before 1980, I utilize province-level employment data and calculate the corresponding share of single-daughter families at the province level. I then classify provinces with a share of single-daughter families higher than the mean share as "high share" regions and provinces with a share below the mean as "low share" regions. Figure 6 compares the trend in the average share of primary employment between these two groups of regions. The figure shows that before the OCP relaxation in 1982, both high and low share regions experienced similar trends in primary employment share, with the share of primary employment declining from around 82% to 74% during the period 1970-1982. This evidence supports the validity of the parallel trend assumption.

After the 1982 OCP relaxation, the trends in primary employment share diverge between regions with high and low shares of single-daughter families. In the first few years from 1982 to 1990, high share regions experience a slightly higher share of primary employment compared to low share regions. However, during the period 1990-2000, high share regions undergo a larger decrease in primary employment share compared to low share regions, which offsets the initial slightly higher share. This indicates that the OCP relaxation may have had a varying impact on the structural transformation of the labour market across regions with different shares of single-daughter families.

4.4 Baseline Results

Table 1. presents the results of estimating equation (1) for the periods 1982-2000 and 1982-1990. The dependent variables are the log of non-agricultural employment and the log of agricultural employment at the prefecture level. The first two columns show the long-run effects using data from 1982 to 2000. The results indicate that regions with a higher share of single-daughter families experienced a larger increase in non-agricultural employment after the OCP relaxation, while agricultural employment did not show a significant response to the policy shock. Specifically, for regions more affected by the OCP relaxation (with a standard deviation of the single-daughter families share of 0.018), the increase in non-agricultural employment was approximately 10.1 percentage points (0.018 * 5.622). However, when examining the short-run results using data from 1982 to 1990 in columns 3 and 4, the pattern reverses. The higher share of single-daughter families in regions is associated with lower non-agricultural employment during this period. The results are robust after controlling for year and prefecture fixed effects.

Indeed, it is important to interpret the results in light of the short-run and long-run effects. While the magnitude of the short-run effect (-7.826) appears larger than the long-run effect (5.622), it is crucial to consider the net effect over time. The baseline results suggest that the long-run positive effect of the OCP relaxation dominates the short-run negative effect, resulting in an overall net positive effect on industrialization.

The key takeaway from the baseline results is that regions with a larger OCP relaxation shock experienced deindustrialization in the short run, but they eventually bounced back and became more industrialized in the long run. This indicates that there is a trade-off between short-run and long-run effects

Table 1: Baseline Regression Results for 1982-1990 & 1982-2000

	1982-2000		1982-1990	
	log non-agri	log agri	log non-agri	log agri
$post_t \cdot S_r$	5.622**	-3.705	-7.826*	-7.057
	(1.77)	(4.29)	(4.58)	(5.64)
year FE	Y	Y	Y	Y
prefecture FE	Y	Y	Y	Y
Observations	408	408	390	390
R-squared	0.944	0.908	0.909	0.906
Adj R2	0.875	0.794	0.794	0.788

Standard errors clustered at prefecture level. Standard errors are shown in the bracket. Dependent variables are log non-agricultural employment and log agricultural employment during 1982-1990 and 1982-2000 respectively. All regressions use population in initial year as weights.

when implementing similar family planning programs or other shocks to fertility. I will delve into the fundamental drivers that underlie this trade-off in section 4.6.

4.5 Robustness Tests

In this section, I conduct a series of robustness tests to validate the findings from the baseline results.

Non-treated province. I focus on provinces that experienced the OCP relaxation, as discussed in section 2, and exclude provinces with no relaxation or only partial relaxations. This approach ensures that the share of single-daughter families, which serves as the policy exposure measure of OCP relaxation, is only relevant in provinces where these relaxation policies were implemented. By examining provinces without the OCP relaxation, I can assess the validity of using the share of single-daughter families as a measurement of policy exposure. In provinces where the OCP relaxation did not occur, the outcome variable should not respond to changes in this share.

Table 2 presents the regression results using the same baseline regression as equation (1) for provinces without the OCP relaxation shock. None of the coefficients are significant at the 5% level, indicating that neither the short-run nor the long-run effect responds to the exposure measurement in regions without the OCP relaxation policy.

Migrants. Another concern is related to the treatment of migrants in the baseline regression. Migrants in China were subject to different policies in some provinces and could have experienced different policy shocks compared to local residents. To ensure the accuracy of the policy exposure measurement, I recompute the share of single-daughter families by excluding all migrants and then re-run the baseline regression.

Table 3 presents the regression results after excluding migrants from the calculation of the share of single-daughter families. The coefficients in this robustness test are very close to the baseline results in

Table 2: Regression Results for Regions without OCP Relaxation

	1982-2000		1982-1990	
	log non-agri	log agri	log non-agri	log agri
$post_t \cdot S_r$	7.176	1.497	-8.782	0.087
	(6.12)	(3.45)	(7.38)	(4.68)
year FE	Y	Y	Y	Y
prefecture FE	Y	Y	Y	Y
Observations	150	150	144	144
R-squared	0.988	0.936	0.941	0.929
Adj R2	0.974	0.854	0.865	0.837

Standard errors clustered at prefecture level. Standard errors are shown in the bracket. Dependent variables are log non-agricultural employment and log agricultural employment during 1982-1990 and 1982-2000 respectively. All regressions use population in initial year as weights.

Table 3: Regression Results without Migrants

	1982-2000		1982-1990	
	log non-agri	log agri	log non-agri	log agri
$post_t \cdot S_r$	5.706**	-3.505	-7.724*	-6.057
	(1.75)	(5.29)	(4.63)	(6.12)
year FE	Y	Y	Y	Y
prefecture FE	Y	Y	Y	Y
Observations	408	408	390	390
R-squared	0.954	0.910	0.919	0.901
Adj R2	0.869	0.797	0.795	0.787

Standard errors clustered at prefecture level. Standard errors are shown in the bracket. Dependent variables are log non-agricultural employment and log agricultural employment during 1982-1990 and 1982-2000 respectively. All regressions use population in initial year as weights.

terms of both magnitude and significance level. This indicates that the exclusion of migrants from the calculation of the share of single-daughter families has little impact on the estimated policy exposure. It is worth noting that migrants only accounted for around 1% of the total working-age population in 1982, so it is not surprising that their impact on the regional OCP exposure during the early stage of the OCP is likely limited.

Rural and urban. In addition to the different treatment towards migrants, the differential treatment of individuals with agricultural Hukou (rural Hukou) and non-agricultural Hukou (urban Hukou) under the family planning policies can also introduce potential biases in the estimation of policy effects. To address this concern, I calculate the policy exposure for the baseline regression using individuals with rural Hukou, as the OCP relaxation only applies to rural Hukou individuals. This allows me to use non-treated urban individuals for a placebo test. If the baseline effect is driven by region-specific factors that affect all

Table 4: Regression Results using Urban Exposure

	1982-2000		1982-1990	
	log non-agri	log agri	log non-agri	log agri
$post_t \cdot S_r$	5.054	-1.090	3.397	-1.502
	(3.90)	(2.67)	(4.45)	(3.30)
year FE	Y	Y	Y	Y
prefecture FE	Y	Y	Y	Y
Observations	408	408	390	390
R-squared	0.945	0.908	0.908	0.905
Adj R2	0.876	0.794	0.792	0.786

Standard errors clustered at prefecture level. Standard errors are shown in the bracket. Dependent variables are log non-agrifacturing employment and log agricultural employment during 1982-1990 and 1982-2000 respectively. All regressions use population in initial year as weights.

population in the same region other than fertility, using non-treated individuals in the same region should generate a similar pattern.

Table 4 displays the regression results using the policy exposure calculated from the share of single-daughter families for urban Hukou individuals. The coefficients are not significant at the 5% level. This finding further supports the validity of the baseline results and indicates that the observed policy effects are primarily driven by the relaxation of the One Child Policy for rural population rather than being influenced by other region-specific factors.

Ethnic minorities. Another factor to consider is the treatment of ethnic minorities. The one-child policy primarily applied to the Han ethnic majority, and there may have been different policy provisions for ethnic minorities. This disparity could potentially introduce bias when estimating the policy effects. To address this concern, I conducted robustness tests in which I restricted the sample to only include the Han ethnic group, excluding ethnic minorities' impact on the policy effects. The results of these tests are presented in Table 5, and the coefficients are close to the baseline regression results. This suggests that the policy effects are consistent and robust when focusing solely on the Han ethnic group.

4.6 Mechanism

In this section, I explore potential mechanisms through which the family planning program affects structural transformation. To be considered as a potential channel underlying the baseline findings, any factor should meet the following criteria:

(i) It should have a direct effect on structural transformation, meaning that changes in this factor can influence the shift from agricultural sector to non-agricultural sector in the economy.

(ii) It should be correlated with the share of single-daughter families, so that the OCP relaxation could lead to changes in this factor, which in turn affects structural transformation.

Table 5: Regression Results without Ethnic Minorities

	1982-2000		1982-1990	
	log non-agri	log agri	log non-agri	log agri
$post_t \cdot S_r$	5.757*	-2.77	-6.087*	-4.957
	(3.06)	(4.19)	(4.55)	(5.505)
year FE	Y	Y	Y	Y
prefecture FE	Y	Y	Y	Y
Observations	408	408	390	390
R-squared	0.936	0.901	0.902	0.899
Adj R2	0.817	0.794	0.812	0.832

Standard errors clustered at prefecture level. Standard errors are shown in the bracket. Dependent variables are log non-agrifacturing employment and log agricultural employment during 1982-1990 and 1982-2000 respectively. All regressions use population in initial year as weights.

(iii) The pattern of this factor should change in both the short run and the long run to account for the observed reverse pattern in the reduced form.

While several factors satisfy the first two criteria, not all of them may fulfill the third criterion. Therefore, I focus on empirically investigating the following promising channels, which have consistently been emphasized in the literature for their potential to meet the first two criteria and check if they can explain the reverse trend together since it is also possible that factors with different dynamics combine to generate the aggregate reverse trend:

Total labour Force: Changes in fertility resulting from family planning policies can directly impact the total labour force in the long run. Moreover, changes in the dependency ratio due to changes in fertility can also affect the labour force in the short run, influencing the structural transformation process.

Quantity-Quality trade-off: Family planning policies may influence investment in human capital and the composition of the labour force in terms of skills and education levels, which in turn affect the skill distribution of the workforce and shape the structural transformation.

Sex Ratio Distortions: The OCP might have led to sex-selective practices, resulting in imbalanced gender ratios. These imbalances can potentially affect gender-specific employment patterns and the allocation of workers across different sectors, thereby influencing the overall structural transformation.

4.6.1 Total labour Force

The total labour force is a critical factor that directly reflects the effect of fertility changes on the economy's workforce. The OCP relaxation, can influence fertility rates directly, leading to changes in the size of the labour force. With fewer restrictions on childbirth, regions experiencing OCP relaxation might have higher fertility rates, resulting in a larger labour force in the long term.

Furthermore, changes in fertility can also impact the labour force in the short run through alterations in the dependency ratio. As the OCP relaxation allows families to have more children, there might be an initial increase in the number of dependent children in the population. This could potentially reduce the labour force participation rate and have implications for the structural transformation process in the short run.

To provide evidence on the reverse trend of total labour force affected by the OCP relaxation shock, I firstly examine the effect of the shock on fertility changes during the period 1982-1990 by regressing fertility on the share of single-daughter families. Then I explore the trend of the total labour force by replacing the outcome variable in the baseline regression with the log total labour force. The regression results are presented in Table 6. Columns 1 and 2 present the regression results on fertility changes for treated provinces (with OCP relaxation) and non-treated provinces (without OCP relaxation), respectively. The result reveals that the OCP relaxation had a significant impact on fertility in treated provinces, leading to an increase in fertility during the period 1982-1990. However, in non-treated provinces, there was no significant change in fertility with respect to single-daughter family shares. Columns 3 and 4 display the results for the total labour force in treated provinces. The findings show that in the long run, there was a significant increase in the total labour force in 2000, possibly as a consequence of the increase in fertility following the OCP relaxation. However, in the short run, there was a decrease in the total labour force in 1990.

It is not surprising that labour force increase in the long run in response to a increase in the fertility during 1982-1990. To further investigate the decrease in the labour force following the OCP relaxation, I use individual-level data to compare the behavior of single-daughter families with other families. Specifically, I run another regression, where the dependent variable is labour force participation, taking a value of 1 if they are working in any sectors, and the main independent variable of interest is the sex of the first child born in 1982, taking a value of 1 only if the household had only one child in 1982, and that child is a girl.

The results are presented in Table 8. Column 1 and 2 show the results for treated and non-treated provinces, respectively, focusing on rural individuals. The findings indicate that in treated provinces, the labour force participation rate decreases by 10.5% in response to having a girl as the first child. However, this effect is not significant in non-treated provinces. These results provide additional evidence supporting the short-run decrease in the labour force is indeed driven by families with more child. i.e., an increase in the dependency ratio decreases the labour participation rate of parents in the short run.

After establishing the reverse trend of total labour force, Next I turn to examine if this explains the reverse trend in structural transformation. To do that, I include the interaction term of the total labour force in the baseline regression. Table 7 presents the regression results with the inclusion of the interaction term of the total labour force in the baseline regression. Notably, the coefficient on the labour force interaction term is significant in both periods. Moreover, the coefficient on the policy exposure term becomes insignificant when the labour force interaction term is included, indicating that the total labour

Table 6: Regression Results for Fertility and Labour Force

	Fertility (treated)	Fertility (non treated)	labour 82-20	labour 82-90
S_r	0.465***	0.248		
	(0.074)	(0.151)		
$post_t \cdot S_r$			6.421***	-8.846*
			(1.124)	(4.809)
year FE	N	N	Y	Y
prefecture FE	N	N	Y	Y
Observations	615	210	408	390
R-squared	0.012	0.010	0.911	0.936
Adj R2	0.010	0.08	0.801	0.835

Standard errors clustered at prefecture level. Standard errors are shown in the bracket.

force is a significant channel through which the policy affects structural transformation.

Indeed, the changes in the labour force can have significant implications for structural transformation through various mechanisms. One such mechanism is the differential factor intensity between the two sectors, where agricultural sector is more land-intensive and non-agricultural sector is more labour-intensive during that period in China. As the total labour force increases due to changes in fertility, it can push more labour to enter the non-agricultural sector compared to the agricultural sector, conditional on a constant relative price. If this holds true, then all families within the same region would have a similar probability of entering the non-agricultural sector and agricultural sector within the same region. structural transformation may not solely be driven by the total labour supply in a region. Another potential reason is that families with more children may make different decisions regarding labour force participation compared to other families, irrespective of the overall labour supply in the region. These within-family decisions can also play a significant role in influencing employment patterns and sectoral allocation, further shaping the structural transformation. To check which channel is the case, I employ the individual level data and regress the sector choice of parents on the the gender of first birth in 1982 with different fixed effects. results are shown in column 3-10 in Table 8.

In columns 3 and 4 of Table 8, the results show that the coefficient for the variable Sex (indicating having a female first child) is only statistically significant for the non-agricultural sector when prefecture fixed effects are not controlled for, as seen in column 4. However, after including the prefecture fixed effects in column 3, the difference in the probability of working in both sectors becomes non-significant. This indicates that the variation in non-agricultural employment is not primarily driven by differences between families within the same prefecture, but rather across different prefectures. Similar results are observed for the period 1982-2000, where columns 7 and 8 show the results for the non-agricultural sector,

Table 7: Regression Results with Labour Force

	1982-2000		1982-1990	
	log non-agri	log agri	log non-agri	log agri
$post_t \cdot S_r$	1.122	-2.704	-1.126	-4.24
	(3.76)	(3.54)	(2.12)	(6.12)
$post_t \cdot Labour_r$	0.021***	0.04	0.018***	0.04
	(0.01)	(5.29)	(0.01)	(4.57)
year FE	Y	Y	Y	Y
prefecture FE	Y	Y	Y	Y
Observations	408	408	390	390
R-squared	0.989	0.912	0.942	0.948
Adj R2	0.984	0.864	0.912	0.943

Standard errors clustered at prefecture level. Standard errors are shown in the bracket. Dependent variables are log non-agrifacturing employment and log agricultural employment during 1982-1990 and 1982-2000 respectively. All regressions use population in initial year as weights.

Table 8: Individual level Regression Results

			1982-1990				1982-2000			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sex_i	-0.105***	0.152	-0.05	-0.09*	-0.42	-0.57	0.97	0.17*	-0.35	0.57
	(0.015)	(0.151)	(0.071)	(0.079)	(0.981)	(1.247)	(1.87)	(0.092)	(0.998)	(1.754)
prefecture FE	N	N	Y	N	Y	N	Y	N	Y	N
Observations	3,321,125	2,351,698	3,321,125	3,321,125	3,321,125	3,321,125	3,321,125	3,321,125	3,321,125	3,321,125
R-squared	0.064	0.054	0.911	0.124	0.921	0.111	0.941	0.124	0.901	0.145
Adj R2	0.054	0.052	0.801	0.075	0.901	0.099	0.924	0.073	0.899	0.137

Standard errors clustered at prefecture level. Standard errors are shown in the bracket.

and columns 9 and 10 show the results for the agricultural sector. This implies that changes in labour force composition, particularly in the non-agricultural sector, are likely to be attributed to variations in labour supply through factor intensity at the regional level rather than household-level decisions based on the gender of the first child.

To provide more evidence on the factor intensity channel, I explore the subsectors within the non-agricultural sector with different levels of labour intensity and run the baseline regression separately based on the labour intensity [6]. The results are shown in Table 9, which provide further evidence supporting

[6]High-intensity manufacturing includes textile and apparel manufacturing, footwear manufacturing, leather goods manufacturing, furniture manufacturing, paper and paper products manufacturing, printing and related support activities, and food processing industries (e.g., meat and dairy products). Low-intensity manufacturing includes electronic manufacturing, automobile production, machinery manufacturing, chemical manufacturing (excluding certain chemical industries), pharma-

Table 9: Regression Results for Different non-agricultural sectors

	1982-2000		1982-1990	
	High intensity	low intensity	High intensity	low intensity
$post_t \cdot S_r$	7.422***	1.243*	-7.826*	-2.854*
	(1.77)	(1.01)	(6.11)	(2.01)
year FE	Y	Y	Y	Y
prefecture FE	Y	Y	Y	Y
Observations	408	408	390	390
R-squared	0.977	0.894	0.965	0.896
Adj R2	0.943	0.874	0.901	0.879

Standard errors clustered at prefecture level. Standard errors are shown in the bracket.

the factor intensity channel. In both periods (1982-2000 and 1982-1990), the coefficients on the policy exposure term for high-intensity manufacturing subsectors are significantly larger in magnitude compared to those for low-intensity manufacturing subsectors.

Based on the above analysis , it is evident that the labour force channel is a crucial factor explaining the reverse trend in the baseline findings regarding structural transformation.

4.6.2 Other potential Mechanisms

In this section, I firstly investigate the **Quantity-Quality trade-off** channel as a potential mechanism through which the family planning program affects structural transformation. The Quantity-Quality trade-off refers to the idea that changes in family planning policies may influence investment in human capital and the composition of the labour force in terms of skills and education levels. This channel explores the possibility that fertility policies have differential effects on the skill distribution of the workforce, which, in turn, affects the sectoral allocation of workers and the overall structural transformation of the economy.

To examine this channel, I firstly analyze the data on education levels before and after the relaxation of the OCP. Specifically, I investigate whether there were any changes in the proportion of skilled workers (e.g., those with higher education levels) in response to the policy relaxation.

Table 10 presents the regression results for the Quantity-Quality trade-off channel, specifically focusing on the share of skilled labour in the labour force. The dependent variable in column 1 and 3 is the share of skilled labour as a proportion of the total labour force, while the dependent variable in column 4 is the share of skilled labour among those under 19 years old within the labour force. None of the coefficients for the Quantity-Quality trade-off channel are statistically significant, indicating that the policy relaxation

ceutical manufacturing, aerospace and defense equipment manufacturing, electrical equipment manufacturing, and computer and electronic products manufacturing.

Table 10: Regression Results for Skilled Share and Sex Ratio

	1982-2000		1982-1990		
	Total skilled share	Sex ratio	Total skilled share	New skilled share	Sex ratio
$post_t \cdot S_r$	0.01	-0.75***	0.03	0.06	-0.65***
	(0.12)	(0.24)	(0.25)	(0.34)	(0.23)
year FE	Y	Y	Y	Y	Y
prefecture FE	Y	Y	Y	Y	Y
Observations	408	408	390	390	390
R-squared	0.989	0.912	0.942	0.942	0.879
Adj R2	0.978	0.864	0.912	0.912	0.861

Standard errors clustered at prefecture level. Standard errors are shown in the bracket.

did not result in a trade-off between quantity (i.e., total labour force) and quality (i.e., skill levels) of the labour force at least at the regional level in my study. Therefore, this is less likely to be a channel explaining my main findings.

Next, I explore the potential role of **sex ratio distortions** as a channel through which family planning policies might have influenced structural transformation. This may affect through the following channel, as the policy relaxation only happen to families with single-daughter, the sex ratio would be distorted for newborns. If newborns of different genders show varying propensities towards working in non-agricultural sector or agricultural sector, this could significantly influence the overall structural transformation of the economy.

The results in Table 10, columns 2 and 5, show relationship between the sex ratio of newborns and the share of single-daughter families. The dependent variable in column 2 is the sex ratio (male to female) within the age group of 0-9, representing the sex ratio of newborns during the policy relaxation period. On the other hand, the dependent variable in column 5 is the sex ratio within the labour force aged under 19, capturing the sex ratio of individuals who entered the labour force during that period. Both are significant and negative, suggesting that the sex ratio is distorted after the OCP relaxation.

To assess the impact of sex ratio distortions on the process of structural change, I approach the question from two angles. First, I include the sex ratio interaction term in the baseline regression to investigate its aggregate effect at the regional level. Second, I turn to individual-level data from the year 2000 to delve into the propensity of male and female individuals towards different sectors to explore how sex ratio distortions might have impacted labour choices at the micro-level.

Table 11 presents the regression results for the baseline model with the inclusion of sex ratio interactions (columns 1-4). While the coefficients in these regression show some decrease compared to the baseline regression, they are not statistically different. Specifically, the coefficient on the sex ratio interaction term

Table 11: Regression Results for Sex Ratio

	1982-2000		1982-1990		2000	
	log non-agri	log agri	log non-agri	log agri	non-agri	agri
$post_t \cdot S_r$	5.324**	-3.354	-7.426*	-5.054		
	(1.89)	(5.14)	(4.76)	(7.21)		
$post_t \cdot Sex_r$	3.45	4.15	6.47	5.81		
	(4.56)	(6.34)	(7.13)	(7.36)		
$Female_i$					0.13	-0.21
					(1.24)	(0.86)
year FE	Y	Y	Y	Y	N	N
prefecture FE	Y	Y	Y	Y	Y	Y
Observations	408	408	390	390	3,321,125	3,321,125
R-squared	0.957	0.917	0.909	0.906	0.103	0.153
Adj R2	0.898	0.805	0.794	0.788	0.098	0.143

Standard errors clustered at prefecture level. Standard errors are shown in the bracket.

is also not significant, indicating that sex ratio distortions do not appear to be a key channel explaining the observed trend in structural change. In columns 5 and 6 of Table 11, I investigate the relationship between gender and sectoral employment patterns specifically for the labour force aged below 20. The dependent variables in these regressions are binary variables indicating whether individuals in this age group are employed in the non-agricultural or agricultural sectors, respectively. Coefficient are not significant, which means that among the newborns after the OCP relaxation, there are no significant differences in the propensity to choose different sectors between males and females within the labour force aged below 20. The findings from both the aggregate and individual-level analyses suggest that the sex ratio distortions may not be a significant factor influencing the overall process of structural change.

5 Quantitative Analysis

In the quantitative analysis section, I calibrate the model using various data sources and model structures to study the aggregate effects of the OCP and its interactions with other forces in the economy. The calibration process involves estimating the values of various parameters in the model based on available data and model equations. Once the model is calibrated, I perform counterfactual analyses to explore the implications of different scenarios including changes to the OCP, as well as other fundamentals in the economy.

5.1 Calibration

In this section, I present the methodology employed to calibrate the model for each parameter. The region in the model is defined to encompass 285 prefectures, and panel data from these prefectures is collected for calibration purposes. However, as data availability may vary, not all variables might be directly accessible at the prefecture level. In such cases, I use data from higher administrative levels, such as province-level data, and make reasonable assumptions to ensure the model accurately matches the observed data.

I estimate a set of structural parameters and fundamentals using model implies equation, reduced form regressions from structural equation, model inversion and simulated method of moments (SMM). Firstly production shares (γ_s^A, γ_s^M), Hukou costs ($\delta_{j,h}^n$), entry fixed cost (f_e), and the share of time cost for each child (q) are backed out directly from the model equation to match data moments. Then cross-region are cross-sector migration elasticities (ϵ and ζ) are estimated through reduced-form regression, while trade elasticities (θ^A, η) are obtained from existing literature. With these elasticities, region-sector productivity A_n^j is calibrated to match sectoral GDP in each region in the data. Since detailed trade and migration data are not available at prefecture level, I use province level data and make some assumptions to estimate migration costs $\tau_{rn,h}$ and trade costs κ_{rn}^j. regional preference for offspring k_i is calibrated to match the model with the fertility rate in each region before1982 [7]. OCP fine $\chi_{n,h}$ is calibrated to match the fertility after 1982. Fertility elasticity η_f is estimated using SMM to match the model predicted elasticity with data. Consumption parameter η_c, ν are backed out using SMM to match aggregate GDP in two sectors and relative price ratios in three years. Table 12 summarises the moments and data used in estimation.

Production shares (γ_s^A, γ_s^M) and fixed entry cost (f_e^n). Because of the structural of Cobb-Douglas production function, production shares is express as the share of input cost as a share of total output. I get these share from the 2002 national IO table [8]. From the free entry condition for non-agricultural sector, the fixed entry cost (f_e^n) cost can be express as a function of only observable and other model parameters, which can be estimated separately: $f_e^n = \frac{L_n^M}{(1+\gamma_L^M(\eta^j-1))N_n^M}$. This equation shows that the entry cost is inversely related to the number of firms in the region. As the number of firms increases, the fixed entry cost decreases, given the same amount of labour supply. On average, the entry cost decrease by 20% during 1982-2000 (from 2.24 to 1.79), in terms of efficient labour requirement.

Hukou costs ($\delta_{j,h}^n$) . The Hukou cost in the model represents the frictions associated with changing sectors, which are distinct from an individual's Hukou type. By manipulating the sectoral choice share

[7] Here I match fetility rate in 1970 in the sense that fertility may already affected by the OCP during 1970-1982(Chen and Fang, 2021)

[8] Steps for calculating these shares are the same as in Tombe and Zhu [2019]. Since there is no capital and sectoral linkage in my model, I attribute the share of capital to land and labour proportionally to make sure they add up to one

$s_{n,h}^{j}$, I derive an expression for the relative Hukou restriction as $\frac{s_{n,A}^{A}}{s_{n,A}^{M}} / \frac{s_{n,M}^{A}}{s_{n,M}^{M}} = (\frac{\delta_{M,A}^{n}}{\delta_{A,M}^{n}})^{\zeta}$[9]. Where $\delta_{M,A}^{n}$ is the Hukou friction for agricultural Hukou to work in non-agricultural sector and $\delta_{A,M}^{n}$ is the Hukou friction for non-agricultural people to work in agricultural sector[10]. If this ratio is large, it indicates that the Hukou system imposes greater restrictions on rural Hukou people to work in the non-agricultural sector. The intuition for this equation is as follows: the share of employment in each sector is determined by two relative forces—sectoral wage and Hukou restrictions. Individuals with different Hukou types face the same wage efficiency rates for a given sector within a region, dividing the relative shares of employment in agricultural and non-agricultural sectors effectively cancels out the influence of relative sectoral wages. As a result, the remaining impact is solely attributed to Hukou restrictions. Figure 7 displays the distribution of the log relative Hukou cost during the period 1990-2005, revealing a consistent decrease in Hukou costs over time. This decline can be attributed to the several Hukou reforms implemented after the 1990s. On average, relative Hukou restriction decrease by 68%.

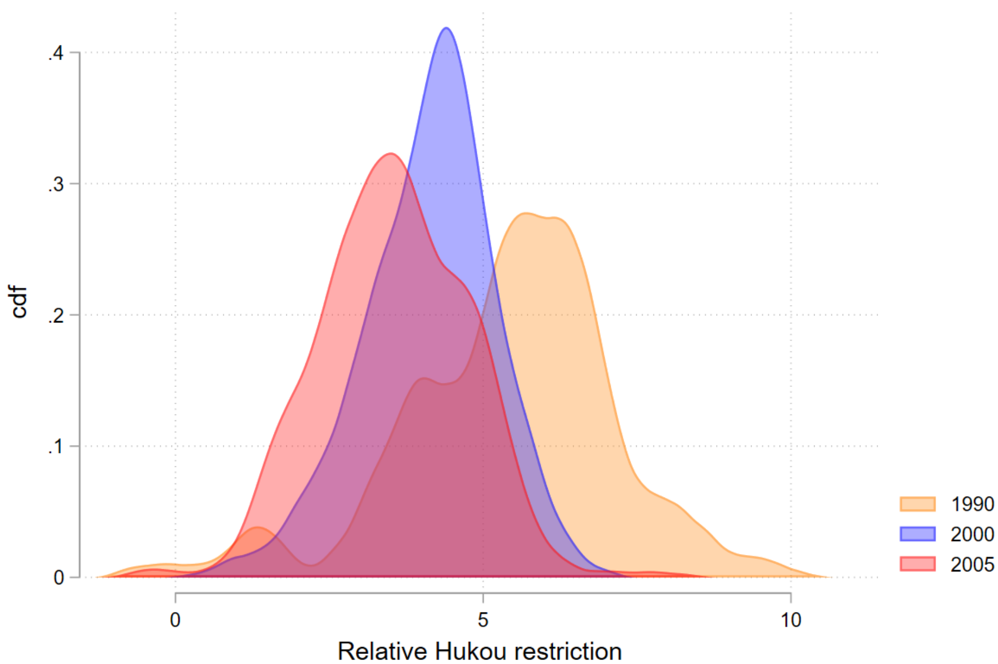

Figure 7: Relative Hukou Restrictions

Cross-sector migration elasticities ζ . I calculate the relative Hukou frictions using observed

[9]I do not estimate the Hukou cost in levels but only relative terms because the Hukou restriction in the same region can only be estimated up to a scale when there exist also regional level migration cost. A simultaneous increase in Hukou cost in the same region is equivalent to an increase in the migration cost to that region.

[10]Generally this no friction for non-agricultural people to work in either sector arising from Hukou system. I keep this symmetric form to maintain the generality of the model.

data on the share of employment in each sector for each Hukou after getting the elasticity from the following regression. By taking log of the relative share expression, the model implies the reduced form type regression:

$$\ln \frac{s_{n,h}^A}{s_{n,h}^M} = \beta_0 + \zeta \ln \frac{\delta_{M,A}^n}{\delta_{A,M}^n} + \zeta \ln \frac{w_n^A}{w_n^M} + \epsilon_{n,h}$$

where I express the relative share of employment in each region as a function of relative Hukou frictions and relative wages. Using relative wage data and employment data, I empirical run this regression, where I use the wage data for previous year as instrument. This gives me an estimates of 1.55, which is close to 1.5 in Tombe and Zhu [2019].

Migration cost $\tau_{rn,h}$ and cross-region migration elasticity ϵ. I assume that the migration cost takes form of quasi-symmetric structure $\tau_{in,h} = \tau_{i,h}^A \tau_{n,h}^B d_{in}^{\eta_\tau}$, where $\tau_{nn,h} = 1$ (Allen et al., 2020). The origin component $\tau_{i,h}^A$ captures the friction faced by individuals when migrating out of region i, reflecting the cost of forgoing the benefits of their local Hukou. The destination component $\tau_{n,h}^B$ measures the friction experienced by individuals when migrating to region i, representing the cost of working in a region with no local Hukou. d_{in} is the distance between region i and region n . η_τ measures the elasticity of migration to distance. Plugging in the functional form of migration cost into the share of migration yields the following reduced-form regression equation:

$$\ln m_{rn,h} = -\epsilon \eta_\tau \ln d_{rn} - \epsilon \ln \frac{\bar{I}_{n,h}}{P_n} + \alpha_r + \epsilon$$

Using migration data, distance and instrumenting the expected real wage with past years' wage, I get the estimation for these elasticities $\epsilon = 1.5$, $\epsilon \eta_\tau = 0.8$. The functional form of migration cost reduces the degree of freedom of the migration cost from $N * N * 2$ to $2N + 2N$, which means that the bilateral migration data is not necessary for the estimation. I calibrate the origin component $\tau_{i,h}^A$ and destination component $\tau_{n,h}^B$ by matching the share of migrants who migrate in the destination and the share of migrants who migrate out of the origin. The migration costs component are only identified up to a scale: any value satisfy $\kappa_i^A = t\tilde{\kappa}_i^A$, $\kappa_i^A = \tilde{\kappa}_i^A / t$ will match the moment in the data. My estimation shows that migration cost is 2 times higher for agricultural Hukou than non-agricultural Hukou on average.

Trade cost κ_{rn}^j. The trade cost κ_{in}^j is assumed to follow a power function of distance, given by $\kappa_{in}^j = \kappa^j d_{in}^{\eta_\kappa}$, where $\kappa_{nn} = 1$. This formulation implies that trade costs increase with distance between regions, and it allows for different trade costs for trade between regions and sectors. η_κ is the trade elasticity to distance and κ^j is the overall level of trade cost in sector j. I assume that η_κ is the same across sectors. The difference in the trade cost is driven by the parameter κ^j. The distance elasticity η_κ can be estimated from the trade gravity equation derived from the model:

$$\ln \pi_{jrn} = -\eta_\kappa \theta_a \ln d_{rn} + \alpha_r + \alpha_n + \epsilon$$

Bilateral trade data is not available at the prefecture level. I use province level bilateral railway trade data to estimate the elasticity of trade to distance. To calibrate the overall level of trade cost, I match the model with total trade share for each sector[11]. The overall level of trade cost decrease by 13% in my estimation.

OCP fine $\chi_{n,h}$, fertility elasticity η_f and regional preference for offspring k_i. I choose values for the regional preference for offspring k_i to precisely match the regional fertility rate before the implementation of the OCP. By doing so, the model can account for the influence of all other factors that affect fertility rates and are consistent over time without being affected by OCP. To capture the effect of the OCP itself, I chooses the OCP fine $\chi_{n,h}$ in such a way that it exactly matches the regional fertility rate after the implementation of the policy taking the value of k_i into account. For the fertility elasticity η_f, I simulate the model with different values of η_f and get the coefficient from regressing fertility rate on regional real wage. Then I choose the value of η_f that best aligns with the observed relationship between fertility and regional real wages. The estimated fertility elasticity parameter η is 0.75, which demonstrate an decreasing return to fertility. The coefficient generated by the model is -0.234, while the corresponding coefficient obtained from the real data is -0.221. This coefficient shows that my model generate a negative relationship between fertility and income without incorporating quantity-quality trade-off. which is widely documented in the literature and often explained by quantity-quality trade-off (Doepke et al., 2022). The estimated mean of the OCP fine for non-agricultural Hukou holders is approximately 35% of their wage, while for agricultural Hukou holders, it is about 23% of their wage. These values are higher than the figures typically found in policy documents, which usually range from 10% to 15%. This difference can be explained by the fact that my estimation considers the monetary equivalent value for all kinds of punishments, including the possibility of losing one's job or facing other non-monetary penalties.

Productivity A_n^j and consumption parameter η_c, ν. Region-sector specific productivity is calibrated by selecting values that precisely align the model with sectoral GDP in each region. To compare productivity across time, I also ensure that the model generated aggregate price index over the years match the Consumer Price Index (CPI) data. To obtain the consumption parameters η_c and ν, I match the model with the aggregate share of GDP for both sectors across different years and also with the relative price index between the two sectors during these periods. On average, productivity increase by 8% and 5% each year for agricultural and non-agricultural sectors.

[11] In 1982, I do not have total trade share but only railway trade share. Therefore I use the relative change in the share of trade by railway from 1985-2000 as the relative change in total trade value to backout the trade value in 1982

Table 12: Parameters and Moments

Parameter	Notation	Moment	Data	Dim
Production shares	γ_s^A, γ_s^M	Cost share in production	IO table	2
Migration cost	$\tau_{rn,h}$	Migration shares	Migration	$N*N*2$
Trade cost	κ_{rn}^j	Trade shares	Trade flow	$N*N*2$
Migration elasticity	ζ	Migration equation	Wage and employment	1
Trade elasticity	θ^A, η	Literature	6	2
Productivity	A_n^j	Sectoral GDP and price	GDP+aggregate price	$N*2$
Hukou cost	$\delta_{j,h}^n$	Employment share by Hukou	Employment	$N*2$
Entry fixed cost	f_e	No.of firms	No. of firm	N
OCP	$\chi_{n,h}$	Mean fertility	Fertility	$N*2$
Consumption para	η_c, ν	Aggregate share of exp.	Expenditure share	2
Fertility para	η_f, q	Fertility elasticity, mean exp.	Fertility and wage	2
Relative preference	k_i	Fertility before OCP	1970 Fertility	1

5.2 OCP Effect

After calibrating the model, I use the model to examine the short-run and long-run effects of the OCP. To begin with, I present the trend between the estimated OCP fine and the fertility rate before and after removing the OCP fine. Figure 8 illustrates the relationship between the estimated OCP fine (in log) and the observed fertility rate (in percentage point) in 1980. As expected, regions with high OCP fines experience lower fertility rates, indicating the influence of the policy on fertility choices. In Figure 9, I plot the changes in fertility if I remove the OCP fine. This is calculated as the fertility rate in a world without OCP fine minus the observed fertility rate. This difference represents the potential gains in fertility that regions missed out on due to the OCP fine. By doing so, I can observe the impact of the OCP fine on fertility choices more clearly, as it isolates the effect of the policy itself from other factors influencing fertility by only changing the fine rate. Regions with higher OCP fines have larger differences, indicating that they would have had higher fertility rates in the absence of the OCP fine.

Because of the structure of the model, higher fertility leads to a greater need for child care inputs in the first period, and these children only become adults in the second period. As a result, in the short run, the labour supply decreases mechanically due to the higher fertility. The long-run labour supply increase when these newborns enter the labour market in the presence of high migration cost, which means a large share of new generation would stay in the same region as opposed to migrate to other regions. The model structure and high migration costs gives the reverse trend of total labour supply following the shock on fertility, in this case, the OCP.

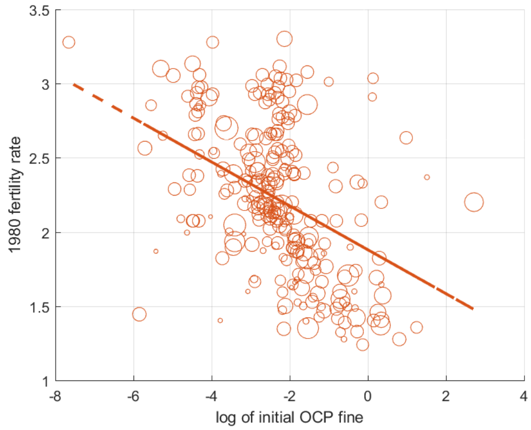

Figure 8: OCP fine vs Observed Fertility.

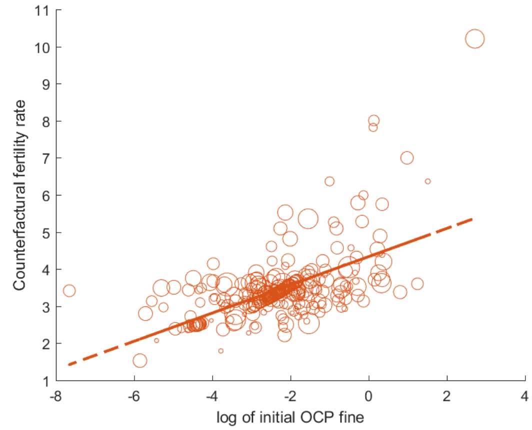

Figure 9: OCP fine vs Fertility Changes

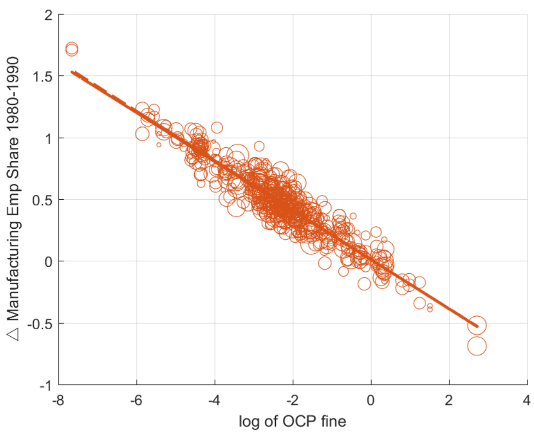

Figure 10: Short run effect

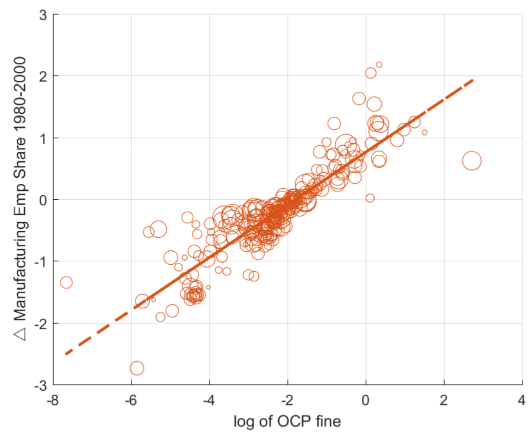

Figure 11: Long run effect

Then I turn to check the response of non-agricultural employment share in 1990 and 2000 following the removal of OCP. Figure 10 illustrates the changes in non-agricultural employment share in the short run with the initial OCP fine. Specifically, I plot the non-agricultural employment share if there were no OCP minus the observed non-agricultural share. The plot reveals a negative relationship between the two, meaning that the increase in fertility following the removal of OCP led to a decrease in non-agricultural employment share. However, in the long run, as shown in Figure 11, a positive relationship emerges, indicating that the non-agricultural employment share rebounds and becomes higher than before in regions with higher initial fertility rates.

To this end, my model generates the same pattern as shown in the empirical findings. The immediate increase in fertility leads to a decrease in labour supply in the short run due to the higher demand for child care inputs. However, in the long run, the positive relationship between fertility and non-agricultural employment share emerges

5.3 Interaction with Other Forces

During the period under study, China experienced significant changes in various fundamentals, such as productivity, Hukou restrictions, migration, and trade cost reduction. These changes could have influenced China's economic landscape and contributed to its structural transformation. In this section, I how the OCP interacts with these other forces in shaping China's economic dynamics, with a specific focus on the long-run effects. As the key difference between the long-run and short-run is the total labour supply, any positive effect on labour supply in the long run can be translated into a negative effect in the short run.

OCP and TFP. Techonolgy progress are regarded as the main drivers of structural change. I consider the counterfactual analysis to study the interaction between OCP and TFP. Figure 12 summaries the effect. In the OCP baseline scenario, I maintain the same values as in Figure 11, which captures the long-run effects of the OCP. Then, to isolate the effects of technological progress, I simulate a world without changes in agricultural or non-agricultural Total Factor Productivity (TFP). In this alternative world, I calculate the changes in non-agricultural employment share between the case where the OCP exists and the case where the OCP does not exist. Therefore, the the difference between the OCP baseline and the scenario without changes in TFP shows the influence of TFP on the marginal effect of the OCP in driving structural change. If there is no increase in agricultural TFP, the removal of the OCP fine would have resulted in a larger increase in the non-agricultural employment share. The reverse trend is found for non-agricultural TFP. Specifically, if there is no increase in non-agricultural TFP, the removal of the OCP fine would have resulted in a smaller increase in the non-agricultural employment share.

OCP and frictions. I perform the same counterfactual analysis to examine the interaction between the OCP and other frictions, including migration cost, trade cost, and Hukou restriction. Figure 13 illustrates that if migration costs are as high as they were in 1982, the removal of the OCP would result in larger changes in non-agricultural employment share. This is because higher migration costs make it less likely for individuals to migrate out of their region to seek employment opportunities in other areas. Then local labour force would increase more than the baseline case which result in a larger responce of non-agricultural employment share. However, if trade costs are as high as they were in 1982, non-agricultural employment share would response less to the OCP removal. When trade costs are high, regional relative price is more affected by regional output. The equilibrium non-agricultural goods prices would decrease more when the local labour supply increase, which puts downward pressure on the equilibrium wage for non-agricultural sector. This, in turn, reduces the attractiveness of the non-agricultural sector for workers, leading to a smaller increase in the non-agricultural employment share in the long run. Figure 14 shows that if the Hukou cost is as high as before, the removal of the OCP would result in smaller changes in non-agricultural employment share. Figure 14 illustrates that if the Hukou cost remains as high as before, the removal of the OCP would lead to smaller changes in the non-agricultural employment share. This finding is reasonable,

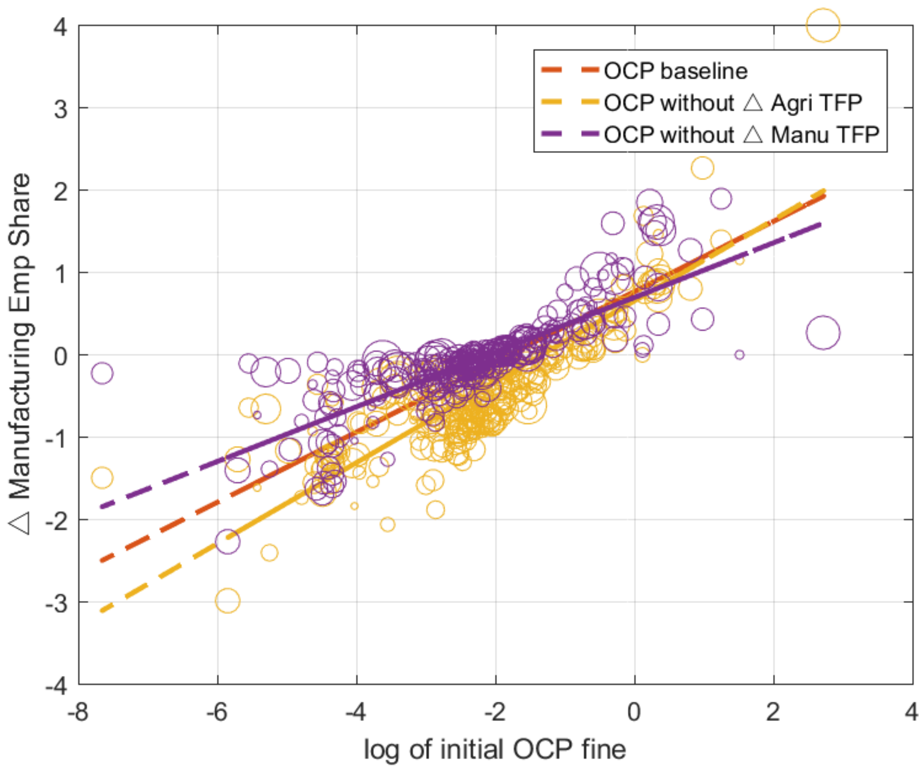

Figure 12: OCP vs TFP

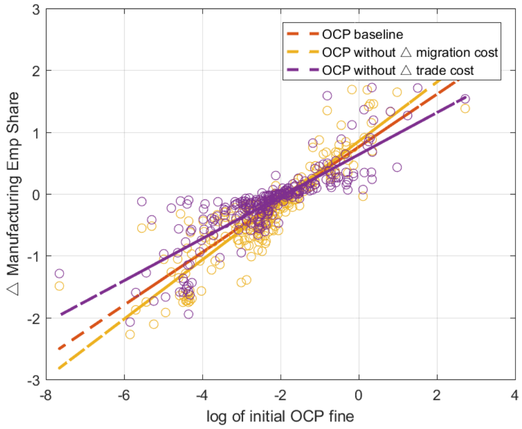

Figure 13: Short run effect

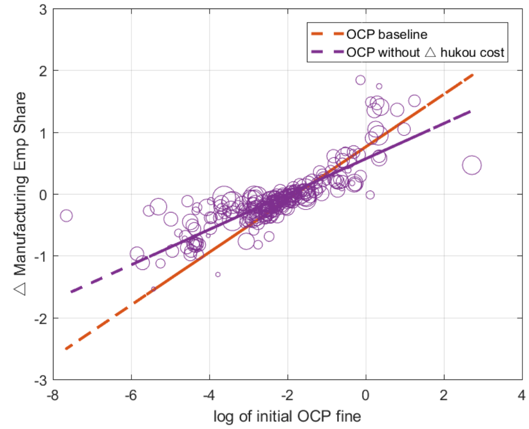

Figure 14: Long run effect

considering that the Hukou cost introduces frictions in labour mobility between sectors. With high Hukou costs, workers are less likely to change sectors, even in response to changes in local labour supply.

5.4 Aggregate Effect

In this section, I assess the aggregate effect of the OCP by comparing the difference between two scenarios: one where the OCP is removed and the other reflecting the real world with the OCP in place.

Table 13 presents the aggregate results for various factors affecting structural change, real wage, and welfare for different Hukou types in the short run during 1982-1990. The percentages indicate the magnitude of changes in the observed outcome variables for each factor compared to the counterfactual scenario where there is no change in that factor during 1982-1990. For OCP result, I compare the real world with OCP with the scenario without OCP. Therefore number in the table shows the effect of each factor on the outcome variables. OCP contributes to a 0.91% increase in structural change in the short run, which aligns with expectations. This magnitude of effect is comparable to the impact of agricultural TFP increase, a traditionally recognized driver of structural change. Additionally, other factors such as increase in non-agricultural TFP and decrease in Hukou costs play significant roles in shaping shaping structural, each contributing to 8% increase in the non-agricultural share. However, the OCP has negative effects on both real wage and welfare. The reduction in fertility following the implementation of the OCP leads to a decrease in the dependency ratio and an increase in the total labour supply. This surge in labour supply exerts downward pressure on real wages. Moreover, the decrease in fertility also decreases utility parents get from children, further contributing to the worsening of welfare. The welfare of non-agricultural Hukou (non-agricultural Hukou) decrease much more than agricultural Hukou as they suffer from higher OCP fines. Increase in agricultural TFP, non-agricultural TFP and decrease in entry cost all contribute to the increase in real wage and welfare for all people. However, the reduction in Hukou cost has large

Table 13: Short Run Aggregate Effect

Factors	structural change	real wage		welfare	
		Agri Hukou	Non-agri Hukou	Agri Hukou	Non-agri Hukou
OCP	0.91%	-0.73%	-1.84%	-4.31%	-14.38%
Agri TFP	0.75%	8.79%	9.17%	6.14%	7.16%
Non-agri TFP	8.14%	3.16%	4.34%	3.17%	4.14%
Entry cost	0.34%	0.01%	2.01%	0.01%	2.19%
Hukou cost	8.13%	2.31%	-0.31%	3.34%	-0.43%

Table 14: Long Run Aggregate Effect

Factors	structural change	real wage		welfare	
		Agri Hukou	Non-agri Hukou	Agri Hukou	Non-agri Hukou
OCP	-2.72%	1.82%	2.44%	-3.24%	-12.44%
Agri TFP	1.12%	16.62%	17.53%	15.51%	18.53%
Non-agri TFP	20.66%	9.37%	12.94%	8.25%	11.18%
Entry cost	1.36%	0.04%	5.09%	-0.03%	5.09%
Hukou cost	24.18%	7.88%	-0.62%	7.47%	-0.83%

positive effect for agricultural Hukou but negative effect for non-agricultural Hukou, due the the fact that a reduction in Hukou cost led to more agricultural Hukou entering into non-agricultural sector and competing with existing non-agricultural Hukou labours.

Table 14 present the long run aggregate effect for these factors during 1982-2000. In contrast to the short run effeect, in the long run, OCP lead to a decrease in non-agricultural employment share by 2.72%. This result is driven by the fact that the total labour force decreases over time following the initial implementation of the OCP. Consequently, there is less labour supply in the market, leading to an increase in real wages for both Hukou types. However, despite the positive effect on real wages, welfare remains negative, indicating that the increase in real wages is not sufficient to compensate for the disutility resulting from the higher cost of raising children. The effects of other forces are similar to those in the short run, except that they increase in magnitude in the long run as these factors experienced larger changes in the long run.

6 Conclusion

In this paper, I investigate the relationship between demographic transition and structural transformation in China, focusing on the implications of the OCP on economic development. By examining the effects of the OCP on fertility rates, total labour supply, and industrialization, I aimed to shed light on the complex interplay between demographic changes and economic outcomes. The empirical findings, along with insights from the dynamic economic geography model, contribute to understanding how demographic policies influence structural transformation and can offer valuable guidance for policymakers facing similar challenges worldwide.

My empirical findings reveal intriguing short-run and long-run effects of the OCP relaxation on China's industrialization process. The mechanism analysis highlights the total labour force as the primary driver behind these effects. In the short run, increased fertility rates lead to a higher dependency ratio, reducing the effective total labour force and resulting in a slower structural transformation. However, this trend reverses in the long run as the children born during the period of increased fertility rates eventually enter the labour market, leading to an increase in the total labour supply. Due to the differences in factor intensity between the two sectors, more labour supply enters the labour-intensive non-agricultural sector than the land-intensive agricultural sector in the long run.

In light of the empirical finding and mechanism, I build a dynamic economic geography model that allows me to quantify the aggregate impact of the OCP and other traditional forces on China's industrialization. The model incorporates dynamic fertility decision-making process, multi-region and multi-sector, non-homothetic preferences, technology progress, and China-specific frictions like Hukou costs and non-agricultural entry barriers. With these features, the model captures China's unique demographic and economic characteristics, providing a comprehensive analysis of the OCP's effects. The model successfully explains and reproduces the reverse trend observed in the empirical findings.

The model's ability to replicate the observed reverse trend in the short-run and long-run effects of the OCP on industrialization is a significant contribution to understanding the complexities of demographic and economic transitions. Using the model and the calibrated OCP fines, I conduct counterfactual analyses to investigate the aggregate effect of the OCP on structural change in China. The findings reveal that the OCP had divergent effects on China's industrialization. In the short run, the OCP led to a 0.91% faster growth rate of industrialization, particularly in regions with stricter enforcement of the policy. This faster pace of industrialization during the early years of the OCP was a consequence of the increased fertility rates, leading to a higher dependency ratio and a reduced effective total labour force. However, in the long run, the OCP had an overall negative impact on industrialization. The reduced labour supply due to the OCP eventually led to a 2.72% slower growth rate of industrialization.

References

D. Aaronson, R. Dehejia, A. Jordan, C. Pop-Eleches, C. Samii, and K. Schulze. The effect of fertility on mothers' labor supply over the last two centuries. *The Economic Journal*, 131(633):1–32, 2021.

T. Allen and C. Arkolakis. Trade and the topography of the spatial economy. *Nber Working Papers*, 129(3):1085–1140, 2014.

T. Allen, C. Arkolakis, and Y. Takahashi. Universal gravity. *Journal of Political Economy*, 128(2):393–433, 2020.

M. Bar and O. Leukhina. Demographic transition and industrial revolution: A macroeconomic investigation. *Review of Economic Dynamics*, 13(2):424–451, 2010.

T. Barham, R. Kuhn, B. McCully, and P. Turner. The demographic transition and structural transformation: Evidence from bangladesh. 2023.

R. J. Barro and G. S. Becker. Fertility choice in a model of economic growth. *Econometrica: journal of the Econometric Society*, pages 481–501, 1989.

G. S. Becker. An economic analysis of fertility. In *Demographic and economic change in developed countries*, pages 209–240. Columbia University Press, 1960.

Y. Bian. Chinese social stratification and social mobility. *Annual review of sociology*, 28(1):91–116, 2002.

L. Brandt, C.-T. Hsieh, and X. Zhu. Growth and structural transformation in china. *China's great economic transformation*, pages 683–728, 2008.

F. J. Buera and J. P. Kaboski. The rise of the service economy. *American Economic Review*, 102(6):2540–2569, 2012.

L. Caliendo and F. H. S. Parro. Trade policy. 2021. URL https://api.semanticscholar.org/CorpusID:240779656.

L. Caliendo, M. A. Dvorkin, and F. H. S. Parro. Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Labor: Supply & Demand eJournal*, 2019. URL https://api.semanticscholar.org/CorpusID:29840906.

K. W. Chan. The household registration system and migrant labor in china: Notes on a debate. *Population and development review*, 36(2):357–364, 2010.

Y. Chen and H. Fang. The long-term consequences of china's "later, longer, fewer" campaign in old age. *Journal of Development Economics*, 151:102664, 2021.

S. Coskun and H. Dalgic. The emergence of procyclical fertility: The role of gender differences in employment risk. 2022.

D. d. l. Croix and O. Licandro. The child is father of the man: Implications for the demographic transition. *The Economic Journal*, 123(567):236–261, 2013.

J. Currie and H. Schwandt. Short-and long-term effects of unemployment on fertility. *Proceedings of the National Academy of Sciences*, 111(41):14734–14739, 2014.

M. J. Delventhal, J. Fernández-Villaverde, and N. Guner. Demographic transitions across time and space. Technical report, National Bureau of Economic Research, 2021.

K. Desmet, D. K. Nagy, and E. Rossi-Hansberg. The geography of development. *Journal of Political Economy*, 126: 903 – 983, 2018. URL https://api.semanticscholar.org/CorpusID:477870.

M. Doepke. Child mortality and fertility decline: Does the barro-becker model fit the facts? *Journal of population Economics*, 18(2):337–366, 2005.

M. Doepke and M. Tertilt. Women's liberation: What's in it for men? *The Quarterly Journal of Economics*, 124(4): 1541–1591, 2009.

M. Doepke and M. Tertilt. Families in macroeconomics. In *Handbook of macroeconomics*, volume 2, pages 1789–1891. Elsevier, 2016.

M. Doepke, A. Hannusch, F. Kindermann, and M. Tertilt. The economics of fertility: A new era. Technical report, National Bureau of Economic Research, 2022.

F. Eckert and M. Peters. Spatial structural change. *SSRN Electronic Journal*, 2022a. URL https://api.semanticscholar.org/CorpusID:46908249.

F. Eckert and M. Peters. Spatial structural change. Technical report, National Bureau of Economic Research, 2022b.

B. Erten and J. Leight. Exporting out of agriculture: The impact of wto accession on structural transformation in china. *Review of Economics and Statistics*, 103(2):364–380, 2021.

P. Fajgelbaum and S. J. Redding. Trade, structural transformation, and development: Evidence from argentina 1869–1914. *Journal of political economy*, 130(5):1249–1318, 2022.

T. Fan, M. Peters, and F. Zilibotti. Service-led or service-biased growth? equilibrium development accounting across indian districts. *NBER Working Paper Series*, 2021. URL https://api.semanticscholar.org/CorpusID:233650147.

O. Galor and D. N. Weil. The gender gap, fertility, and growth, 1993.

J. Greenwood, N. Guner, and G. Vandenbroucke. Family economics writ large. *Journal of Economic Literature*, 55 (4):1346–1434, 2017.

J. Greenwood, N. Guner, and R. Marto. The great transition: Kuznets facts for family-economists. Technical report, National Bureau of Economic Research, 2021.

B. Herrendorf, R. Rogerson, and A. Valentinyi. Growth and structural transformation. *Handbook of economic growth*, 2:855–941, 2014.

V. J. Hotz, J. A. Klerman, and R. J. Willis. The economics of fertility in developed countries. *Handbook of population and family economics*, 1:275–347, 1997.

Z. Huang, L. Lin, and J. Zhang. Fertility, child gender, and parental migration decision: Evidence from one child policy in china. *Available at SSRN 3375122*, 2019.

L. E. Jones, A. Schoonbroodt, and M. Tertilt. Fertility theories: can they explain the negative fertility-income relationship? Technical report, National Bureau of Economic Research, 2008.

O. M. Leukhina and S. J. Turnovsky. Population size effects in the structural development of england. *American Economic Journal: Macroeconomics*, 8(3):195–229, 2016a.

O. M. Leukhina and S. J. Turnovsky. Push, pull, and population size effects in structural development: Long-run trade-offs. *Journal of Demographic Economics*, 82(4):423–457, 2016b.

R. E. Manuelli and A. Seshadri. Explaining international fertility differences. *The Quarterly Journal of Economics*, 124(2):771–807, 2009.

O. Moav. Cheap children and the persistence of poverty. *The economic journal*, 115(500):88–110, 2005.

L. R. Ngai and B. Petrongolo. Gender gaps and the rise of the service economy. *American Economic Journal: Macroeconomics*, 9(4):1–44, 2017.

L. R. Ngai and C. A. Pissarides. Structural change in a multisector model of growth. *American economic review*, 97 (1):429–443, 2007.

L. R. Ngai, C. A. Pissarides, and J. Wang. China's mobility barriers and employment allocations. *Journal of the European Economic Association*, 17(5):1617–1653, 2019.

N. Qian. Quantity-quality and the one child policy: The only-child disadvantage in school enrollment in rural china. Technical report, National Bureau of Economic Research, 2009.

S. J. Redding and E. Rossi-Hansberg. Quantitative spatial economics. *Annual Review of Economics*, 9:21–58, 2017.

W. C. Robinson and J. A. Ross. *The global family planning revolution: three decades of population policies and programs*. World Bank Publications, 2007.

M. R. Rosenzweig and J. Zhang. Do population control policies induce more human capital investment? twins, birth weight and china's "one-child" policy. *The Review of Economic Studies*, 76(3):1149–1174, 2009.

T. Schultz. Demand for children in low income countries. chapter 8 in handbook in population and family economics, vol. 1a, edited by mr rosenzweig and o stark, 1997.

D. J. Solinger. *Contesting citizenship in urban China: Peasant migrants, the state, and the logic of the market*. Univ of California Press, 1999.

Q. Song and J. P. Smith. Hukou system, mechanisms, and health stratification across the life course in rural and urban china. *Health & place*, 58:102150, 2019.

T. Święcki. Determinants of structural change. *Review of Economic Dynamics*, 24:95–131, 2017.

T. Tombe and X. Zhu. Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review*, 109(5):1843–1872, 2019.