Explaining Trends in Adult Height in China: 1950 to 1990*

Minhee Chae†[1], Tim Hatton‡[2], and Xin Meng§[3]

[1]Nankai University
[2]Australian National University and University of Essex
[3]Australian National University

July 8, 2022

Abstract

This paper explores the changing trend of adult height in China for cohorts born in 1950-90. We use information on household structure and local economic conditions during the individual's childhood to explain the trend. We find that during the 40-year period, the growth rate of adult height increased, with the most substantial increase occurring in the 1980s. One important contributing factor to the growth of adult height is the continued increase in government per capita spending on health and education. The impressive growth in the 1980s was mainly due to the introduction of market-oriented economic reforms, rather than the advent of the One-Child Policy. We find that the positive effect of economic reforms was larger for urban dwellers than for their rural counterparts and within the rural areas the benefit was far greater for men than for women.

Keywords: Height; China; economic reform

JEL Classification Code: I15; I18; J13; O1

*The authors are grateful for financial support from the Australian Research Council (DP140103603)
†minheechae@nankai.edu.cn
‡tim.hatton@anu.edu.au
§xin.meng@anu.edu.au

1 Introduction

Spectacular development in the countries of East Asia in recent decades has been associated with dramatic improvements in a range of health indicators, including average stature. A recent study puts China at the top of the league table (NCD, 2020). Adult height is a key marker of socioeconomic conditions in the household and the locality during the first few years of life and a large literature assesses these influences (Steckel, 2009; Perkins et al., 2016). While China's tumultuous history from the 1950s to the 1980s and its social and economic consequences are well known, the implications for health and height are less clear. How were those turbulent times manifested in the health and heights of the Chinese population? What local indicators were associated with changes over time and between urban and rural areas? And, above all, how did the economic reforms from the early 1980s accelerate improvement and through which mechanisms? In this paper we chart in detail the heights of individuals born from 1950 to 1990 and we explore the association with a variety of indicators at the household and local level around the time of birth.

Studies that examine the changes in adult height of cohorts born from the 1950s to the 1990s are comparatively rare. In this paper we focus on three questions. First, what are the long term trends over these decades and how do they differ between rural and urban areas and by gender? Second, how has the evolution of economic and social conditions and policies during early childhood been associated with individuals' adult heights? Third, and most important, how did family planning policies and economic reforms, that were introduced at the end of the 1970s differentially affect the height trajectories of these different sectors of the population?

To answer these questions, we use individual adult heights from the China Family Panel Study (CFPS). This survey has two important advantages. One is that it is nationally representative. The other is that it records individuals' childhood information: birth county, household registration (hukou) status at age 3, and some basic parental information at the time the person was aged 14. This allows us to link an individual's adult height to the local socioeconomic conditions that he or she faced during early childhood. To do so we build a unique set of county-level data on economic conditions (industrial and agricultural outputs or total outputs), a general health indicator (mortality rate), and local government expenditure on social services between 1950 and 1990.

Briefly, we find that average heights of rural and urban males born between 1950 and 1990 increased by 1.3cm and 1.5cm per decade respectively, while for females, heights increased at 0.7cm and 1.1cm per decade respectively. Regression estimates reveal that government expenditure on health and education is positively associated with height while the association with measures of industrial and agricultural output is significantly positive only for males, and mainly in rural areas. We also find a distinct acceleration in the height trend in the period of post-1979 economic reforms. Controlling for the effects of economic reforms we find that the negative effect of family size on height was enhanced under the one-child policy (OCP), but that the effect is attenuated for females and in rural areas where son-preference was prevalent.

Our results contribute to several strands of the literature. One is the study of long term trends, which has often been based on less representative samples and/or suffers from gaps in the record. For example, using evidence from industrial surveys Morgan (2004), finds very modest growth in height of cohorts from the 1880s to the 1920s and, using data on school children, Morgan (2000) shows much faster growth after 1979. Our results confirm the 1980s turning point but indicate that the pace of increase in the preceding decades was faster than it had been before 1930. They are broadly consistent with Bi and Ji (2005) who find that the mean heights of 18-year-old students in Shandong increased by 1.75cm and 1.07cm per decade for males and females respectively from 1956 to 2000. Other studies have also identified divergence in height between males and females during this period (Schwekendiek and Baten, 2019) but do not distinguish between those with rural and urban backgrounds.

Historical studies of height for China and elsewhere often use variables observed at the time of the survey rather than during childhood (see, for example, Schwekendiek and Baten, 2019) for China, Perkins et al. (2011); Mamidi et al. (2011) for India, and Subramanian et al. (2011) for 54 developing countries. In contrast, we exploit variables such as household income and family size that refer to the individual's childhood. But most important, we link individuals' adult height to local conditions experienced during childhood, which very few studies have been able to do [see Almond et al. (2019); Bailey et al. (2016)]. We are thus able to assess the health effects of local conditions, notably the disease environment and public health provision, which have been widely recognised (Hipgrave, 2011) but rarely demonstrated.

We also account for the shift in trend in heights due to post-1979 changes, notably the economic reforms and the OCP. A range of studies have used the OCP-induced fall in family size to test the trade-off between child quality and quantity, using indicators of health and education as outcome variables (Qian, 2009; Liu, 2014; Kubo and Chaudhuri, 2017; Liang and Gibson, 2018; Zhang et al., 2020). With the important exception of Weng (2019) these studies rely for variation on post-1979 adjustments to the OCP rather than comparing before and after and they often fail to allow for the economic reforms that occurred at the same time, which have been shown to be important for social and economic outcomes (Lin, 1992; Almond et al., 2019). We find that when the shock of economic reforms is taken into account, height, as predicted by the quantity-quality trade-off theory, is negatively associated with family size, especially in the reform period, but the gains were least for females in rural areas where son-preference was prevalent.

The rest of this paper is organised as follows. Section 2 outlines the background of socioeconomic changes in China between 1950 and 1990, and the details of economic reforms and family planning policies, particularly from 1979. Section 3 presents model specifications and estimation issues, and Section 4 introduces the data and the main sample for our analysis. Section 5 presents the main results and discusses potential channels of influence. Section 6 provides robustness checks. Section 7 concludes.

2 Background

In this section we describe socioeconomic conditions and policy changes that have affected the height of people born in the 40 years from 1950 to 1990.

2.1 A centrally-planned economy and its health care system 1949-1977

When the Communist Party of China (CPC) came to power in 1949, China was suffering from severe poverty and health problems associated with malnutrition. The infant mortality rate ranged from 12% to 20% among live births in rural regions and almost a third died before the age of five (Liang et al., 1973; Salaff, 1973).[1] Although the CPC introduced a planned economy upon establishment of the People's Republic of China (PRC), this coexisted with the free market economy until 1955 (Bettelheim, 1988). In rural areas, land reform ensured that all farmers possessed a small piece of land. During the early years, the GDP growth rate was relatively high, albeit rising from a very low level (NBS, 1999).[2]

The new government introduced a health care system, which emphasised equity on the demand side and the central role of the government on the supply side (Yip and Hsiao, 2015).[3] The urban healthcare system consisted of two main schemes: the Government Insurance Scheme, which mainly

[1] According to Salaff (1973), there was very little difference in infant mortality between rural and urban China at that time.

[2] The annual growth rate of gross domestic product (GDP) per capita was 8.72 percent in China between 1952 and 1955 (NBS, 1999).

[3] See Liang et al. (1973) for more detailed policy information.

covered public servants and university staff, and the Labor Insurance Scheme, which covered employees in state-owned enterprises (Xin et al., 2014). The benefits included comprehensive welfare, such as old-age pensions and maternity benefits, in addition to free medical care (Meng, 2000). As the prices of all goods and services were determined by the central government, the cost of providing public services was kept very low, in line with the low wages received by urban workers. It is important to note, though, that this system only covered urban people. For rural residents (accounting for 80% of the population), health care and all other types of welfare were covered by a cooperative system in the villages or communes (tens of villages) that were established after the collectivisation of agriculture in the late 1950s (Meng, 2000; Liu and Cao, 1992).

From 1956, the Chinese government accelerated the nationalisation of all non-agricultural parts of the economy and the national economic plans concentrated on boosting industry at the expense of agriculture (Chow, 1993). Enterprises were owned by the state, they produced goods in line with the national plan, and all profits belonged to the state (Wang, 2018). In rural areas, the collectivisation movement encouraged individual farmers to pool their land to form production teams, subsequently merged into brigades and communes (Lin, 1988). Under the commune system, all resources were collectively owned. Part of the output was procured by the central government at below market price, and the remainder was distributed to the team households according to their basic needs and household members' annual labour inputs. Revenues from government procurement were reinvested into agricultural production and used to provide communal services (Lin, 1988; Meng, 2000).

Communes and brigades established their own clinics to provide primary health services, which gradually became the cornerstone of the rural healthcare system known as the Cooperative Medical System (CMS) (Zhu et al., 1989; Yip and Hsiao, 2015). Some local farmers received basic medical training and became the main medical service providers in rural China (Zhu et al., 1989; Yip and Hsiao, 2015). By 1958 about 20% of the production brigades had adopted the CMS system (Liu and Cao, 1992). As medical services were provided within the communes and brigades, medical professionals were subject to the same distributional rules as farmers and so medical service prices were kept low (Yip and Hsiao, 2015).

Beginning in 1958, the Great Leap Forward (GLF) induced many farmers to abandon agriculture and focus on primitive steel production while the aggressive push for rural collectivisation dramatically reduced their incentives for agricultural production. These policies, together with the inflexible grain procurement system and many other shocks, resulted in the Great Famine (Lin, 1990; Li and Yang, 2005; Gørgens et al., 2012; Meng et al., 2015). Between 1959 and 1961, malnutrition was widespread, and an estimated 16.5 to 45 million people, mostly living in rural areas, perished from hunger or disease (Ashton et al., 1984; Chen and Zhou, 2007; Li and Yang, 2005; Meng et al., 2015).

After the crisis, the government restored the pre-GLF system for agricultural production (Meng et al., 2015), and attempts were made to improve production efficiency (Lin, 1990). Grain output increased and mortality rates fell back to the pre-famine trend. Expenditure on disaster relief was doubled and social welfare funds increased by 7.1% per annum from 1962 to 1965 (NBS, 1999). This period also saw a substantial increase in the quantity and quality of medical services, mainly focusing on disease prevention. The number of laboratories for medicine and chemical reagent tests increased from 50 in 1958 to 151 in 1965 and the number of specialized prevention stations increased dramatically from 1963 (NBS, 1999). However, most of the government-provided services were located in urban centres. Rural health services, for more than 80% of the population, were left to the collectives to administer (Zhang and Unschuld, 2008).

The recovery did not last long and the Cultural Revolution launched in 1966 aimed to purge traditional 'anti-revolutionary' elements and 'capitalists' inside the Party (MacFarquhar and Schoenhals, 2006; Booth et al., 2018). Schools in urban areas were closed; most urban production and some rural agricultural production was stopped (Booth et al., 2018). The annual growth rate of per capita GDP turned negative in the early, chaotic years of the Cultural Revolution, but production subsequently recovered to the pre-Famine level (Borensztein and Ostry, 1996).

During this period, government investment in health services was limited and the training of medical personnel in urban areas collapsed as medical schools were closed until 1972 (Dong and Phillips, 2008). However, the village- and commune-level healthcare services were formally evolved into the CMS and by the mid-1970s, almost 90% of the rural population participated in the system (Liu and Cao, 1992; Liu, 2004), which provided equal access to inexpensive medical care (Tu, 2016). A nationwide programme of basic training was provided for farmers, who were middle-school graduates from rural schools, and for urban youth sent down to the countryside during the Cultural Revolution (Hesketh and Wei, 1997; Zhang and Unschuld, 2008; Dong and Phillips, 2008; Hipgrave, 2011). The so-called barefoot doctors promoted basic hygiene, preventive healthcare and family planning, and treated common illnesses in the countryside (Gong and Chao, 1982). As a result, the number of rural health workers increased from 1.0 million in 1970 to an estimated peak of 1.8 million in 1977 (Hesketh and Wei, 1997; Hipgrave, 2011). The CMS system is widely regarded as a success and infant mortality dropped from 20% in 1949 to 4.7% in 1975 (Liu et al., 1995).

This period also saw the initiation of family planning policies. In the early 1970s, post-famine population growth and the memory of the hunger during the famine ignited a philosophical debate over the Malthusian Population Trap. The policy of "Later, Longer, and Fewer", introduced at the end of 1973 (Center for Population Studies and of the China's Population Yearbook, 1986; Peng, 1991; Feeney and Wang, 1993; Cai, 2010), encouraged couples to get married later, to have longer birth intervals, and to have fewer children. The policy had a significant impact on the birth rate, especially for the urban population. Over the 1970s, the total fertility rate fell from nearly six to just above two (Wang, 2011; Gietel-Basten et al., 2019).

2.2 Market-oriented economic reforms and the change of the healthcare system

Although efforts were made to improve basic living standards and to reduce inequality, after 30 years of socialist revolution, the system had failed to improve people's living standards substantially. At the end of the 1970s, 40-80% of the rural population in China was still living in poverty (Ravallion and Chen, 2007).[4] In response to the poverty in rural areas, covert privatisation began in villages, whereby farming families cultivated their own contracted land and earned income from what they produced, and this resulted in bumper harvests (Lin, 1992; Meng, 2012). Seeing the positive effect of privatisation, the central government officially introduced this new system - the Household Responsibility System (HRS) - to the whole agricultural sector in the late 1970s and early 1980s.

The persistent problem of food shortages was resolved (Du et al., 2014; Wang, 2018) as the gross value of agricultural output increased by 26% between 1980 and 1983 (at constant 1980 prices) (Lin, 1988). Although privatisation in agriculture successfully incentivised farmers and significantly increased agriculture productivity and rural income, it destroyed the old collective system. As the communes dissolved, so too did the system of social services in rural society. Teachers and healthcare workers in rural areas now lost their income source and were unable to provide basic education and health care. In 1978, 82% of all brigades were under the CMS, but by 1983 only 11% brigades remained in the scheme (Huang, 2011; Bloom and Xingyuan, 1997). With the exception of vaccination and family planning, rural healthcare between the 1980s and the early 2000s relied mostly on private health services (Liu et al., 1995; Blumenthal and Hsiao, 2005).

In the urban economy, the privatisation and marketisation process shifted a large proportion of workers from state-owned to privately-owned work units, which significantly increased urban income. But at the same time the old cradle-to-grave social welfare system, long enjoyed by urban residents in the pre-reform era, gradually eroded.

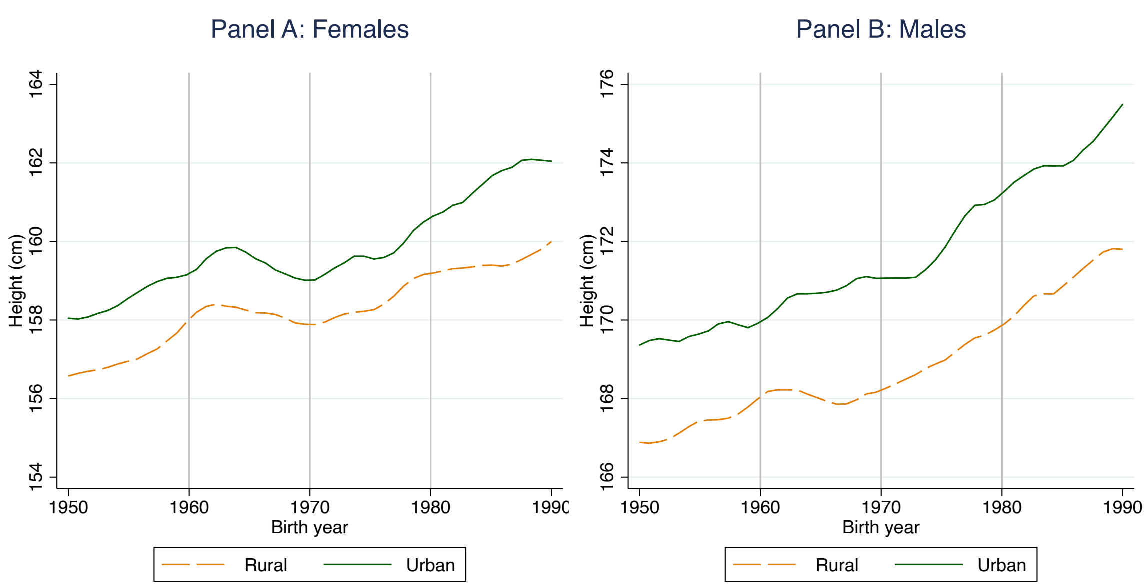
[4]Depending on the poverty line used to measure poverty.

It was not until after the mid-2000s that the establishment of a new social welfare system in both rural and urban areas became one of the priorities of economic reform. To date, the rebuilding of the new rural and urban health care systems are well underway, but their integration will still take a considerable time to achieve.

2.3 The One Child Policy

After the early success in implementing family planning policies (the 'Later, Longer, and Fewer' policies) to bring down fertility, the much stricter 'One Child Policy' (OCP) was put in place in June 1979 . It was rigidly implemented in the urban areas, whereas in rural areas, where cultural traditions have always had a stronger influence, a second birth was allowed if the first child was a girl to accommodate China's 'son-preference' tradition (Peng, 1991). Out-of-quota births were subject to fines, but to different degrees depending on local officials' policy enforcement powers and political lineage (Edlund et al., 2013). Even with a much looser policy, the abnormal sex ratio at birth became a serious issue in rural areas.

Figure 1: Height by rural-urban and gender: 1950-1990



Source: China Family Panel Survey; China Health and Nutritional Survey; Rural-Urban Migration in China
Notes: The graph is generated with local polynomial smoothing line with bandwidth of 1.2.

In general, when fertility decreases, ceteris paribus, family size declines and per-child nutrition intake increases, which should improve health and hence height. However, under a strong culture of son-preference and a potential financial penalty for additional births, it may affect sons and daughters differently. Figure 1 shows average height by rural/urban registration status at age 3 by gender. It indicates that rural men and women are on average shorter than their urban counterparts who were born in the same year, by about 4 cm and 2 cm respectively. However, on trend, rural men track urban men quite closely over the entire period, but rural women's growth in height stalled after the introduction of the OCP (cohorts born after the late 1970s).[5]

3 Model specifications and estimation issues

The unconditional means in Figure 1 may be influenced by compositional changes in population over time. Thus, to answer our first question, and to quantify the trends in height, we estimate a simple regression, controlling for birth-county fixed effects. Specifically, we estimate two different specifications of the following equation:

$$H_{ic} = \alpha_0 + \alpha_1 T_i + \delta_c + \varepsilon_{ic} \tag{1}$$

where H_{ic} is the adult height of individual i born in county c and measured at the survey year. T_i represents the trend in height across birth cohorts, either as a linear trend, or measured as three birth-decade dummy variables (the 1960s, 1970s, and 1980s) with the 1950s used as the omitted category. δ_c is the birth-county fixed effects and ε_{ic} is the random error.

To understand how local economic and social environments affect individuals' stature, we estimate the following equation:

$$H_{ic} = \beta_0 + \beta_1 \sum_{n=1}^{3} \omega_n E_{cn} + \beta_2 X_{i14}^P + \beta_3 W_{ic} + \beta_4 Sib_{ic} + \lambda_c + \epsilon_{ic} \tag{2}$$

where E_{cn} is a vector of socioeconomic conditions in county c at time t_n; n represents the individual's age between 0 and 3 (the first three years of a child's life), while ω_n is a damping coefficient, which varies with age. In this study we focus on the first three years because, according to WHO height-for-age median values for children, height increases most in the first three years of life, accounting for almost 40% of the total height growth up to the age 19 for both boys and girls.[6] ω_n is calculated

[5] The fast catching up of rural female height to that of urban females during the famine years may be related to the survival bias discussed in Gørgens et al. (2012).

[6] Although the first three years only accounts for 40% of the total increase in height, there is a fairly strong correlation between height and health at age three and adult height and health (see, for example, McCormick et al., 2006; Doyle et al., 2009; Cole and Wright, 2011). This suggests that the first three years probably accounts for more

based on the WHO height-for-age median value growth chart. Thus, we use the share of the first three years' height growth as the weights for social economic conditions at the time of the individual's first three years to generate a weighted average of local conditions.[7] In this specification, we replace the birth year trend, T_i, with a vector of local socioeconomic conditions which also vary across birth years to see how the over-time variations in these conditions affected average height. E_{cn} consists of four different measures of socioeconomic conditions at the county level: the log of industrial and agricultural outputs per capita, separately, (In Table A2, we also use the variable 'total output per capita' to test the sensitivity), the log of government social expenditure per capita, and the population mortality rate. X_{i14}^{P} is a vector of parental characteristics at the time the individual was aged 14,[8] including parental education, whether the father was a party member, and the rank of the father's occupation. This parental information captures family income differences over and above those represented by local socioeconomic conditions during the individuals's childhood. W_{ic} is a vector of variables including individual i's ethnicity, gender, and birth order index.[9] Sib_{ic} is the number of siblings individual i has. λ_c is birth-county fixed effects and ϵ_{ic} is the random residual.

In equation (2) we try to mitigate the potential endogeneity problem by using lagged economic and social conditions together with county fixed effects. We argue that, controlling for time-invariant county factors, it is unlikely that individuals' adult height could affect the socioeconomic conditions in their childhood (the reverse causality issue). But there could still be omitted variable problems. For example, parental height is an important determinant of child height, but this could be correlated with local conditions. Although only a very small proportion of the observations contain parental height, including it makes little difference to the results (as noted in Section 6).

Equation (2) assumes that the impact of socioeconomic conditions is constant over the forty-year period. However, during this period China has experienced three important historical events: the

than 40 percent of the variation in final height and health. In fact Cole and Wright (2011) show that the correlation coefficient between height at age 3 and final height is about 0.75 for boys (Figure 1 in Cole and Wright, 2011). In addition to the importance of the first three year of life in height growth, using a larger age window requires much longer time series of the county level data, whereas our data are limited to the years 1950-90. Nevertheless, in Tables A10 and A11, we present the results of sensitivity tests using different age windows: 0 to 5 and 0 to 7. Our results are not sensitive to these changes even though they come at the cost of a loss of observations and a loss of between-cohort variation.

[7] The details of how the weight is calculated are presented in Appendix B.

[8] It would be ideal if we had parental information at a much earlier age, as stature development mainly occurs in the first three years of an individual's life according to the WHO height-for-age median value chart. However, the earliest information we can obtain from the survey is at age 14.

[9] The birth order index is a normalised birth order, which is calculated based on Hatton and Martin (2010). The reason for using the index instead of the actual birth order is because the variables 'number of siblings' and 'birth order' are highly correlated. In our samples, the correlation coefficient between the number of siblings and the actual birth order is 0.6748, while after normalisation the coefficient of correlation reduced to 0.1091.

Great Famine, the Cultural Revolution, and the economic reform after 1978. Despite these changes, the major structural change only occurred after the economic reform when the Chinese economy switched from the planned to the market oriented system — bringing dramatic increases in living standards. Thus, in this study we focus mainly on the potential impact of the economic reform on height. We do, however, test the robustness of our results to the introduction of the other two shocks later in the paper.[10]

To test the changes brought about by the economic reform, we introduce an exposure measure, which indicates what proportion of individual's crucial growth period (age 0-3) occurred during the economic reform period (1979 and after). To do so, we augment equation (2) above as follows:

$$H_{ic} = \gamma_0 + \gamma_1 \sum_{n=1}^{3} \omega_n E_{cn} + \gamma_2 X_{i14}^{P} + \gamma_3 W_{ic} + \gamma_4 Sib_{ic} + \gamma_5 Expo_{ic} + \gamma_6 Expo_{ic} * Sib_{ic} + \mu_c + \upsilon_{ic} \quad (3)$$

where $Expo_{ic}$ measures the share of the individual's first three years of life, weighted by the damping coefficient, which takes a value of zero for those born before 1977, one for those born in 1979 and after, and a fraction between 0 and 1 for those born in 1977 or 1978.[11]

However, there is a complication. The economic reform was introduced in 1979, and so too was the OCP.[12] Thus, the exposure variable for the economic reform can only measure the net effect of the two policies together. Our purpose is to tease out, as much as we can, the pure economic reform effect and to better answer our third question of how the OCP and accompanying economic sanctions for violations of the policy may have affected the two gender groups differentially due to the son-preference culture. To do so we interact the exposure variable for the economic reform ($Expo_{ic}$)

[10]As discussed before, the Great Famine generated significant malnutrition and mortality, especially in rural China. Gørgens et al. (2012) show that because of the high famine mortality rate, which truncated the left tail of the height distribution for the famine cohorts, the survivors are genetically taller individuals. Thus, on average the height difference between the famine survivors and their non-famine counterparts are hard to detect. With regard to the CR, apart from during the first three years when there were some interruptions to the general public services, there is no obvious reason to believe there were any other factors which may affect the height of the rural and urban individuals. If anything, rural public health services were improved during the CR period as we discussed in the background section.

[11]In the sensitivity tests, we also examine whether controlling for the Chinese famine and the Cultural Revolution alter our results. There, we also adjust the exposure to the two events by a damping coefficient. The way these damping coefficients are calculated are the same (see Appendix B).

[12]Note that both economic reform and OCP have different channels through which they can affect height of individuals. Economic reform increased household income, but in its early years, it reduced rural household access to health care. The situation did not improve much until the late 1990s. The net effect of the reform should be a combination of these two potentially opposite effects. Similarly, the introduction of the OCP should reduce sibling numbers, hence increase resources to be shared among siblings within the family. But for households that prefer sons over daughters, the out-of-quota birth in order to get a male birth would increase the number of siblings. What is more, if an out-of-quota birth occurs it would incur a financial fine for the family and further reduce the resources to be shared among siblings.

with the number-of-siblings variable. We argue that the effect of OCP on height mainly goes through fertility. Conditional on household income during childhood (which is captured mainly by variables on parental characteristics at the time the individual was 14 years of age), 'fertility' interacted with the $Expo_{ic}$ has implications for the resources available to our sample individuals, not only through the direct effect of a reduced number of siblings due to the OCP, but also through potential fines incurred for violation of the OCP birth quotas (an additional negative income). Because we control for both the level effects of 'exposure' (post-economic reform positive income effect) and 'number-of-siblings' (pre-OCP fertility effect) variables, the interaction term captures the sibling effect for individuals who were born in the period of post-OCP policy. The coefficient on the un-interacted variable $Expo_{ic}$ captures the economic reform effect for individuals with no siblings, although the effect for any number of siblings can be calculated.[13] We estimate this specification separately for male and female samples to gauge the differential effect of the OCP on the two gender groups.

4 Data and descriptive statistics

4.1 Data

We use two main data sources: the China Family Panel Studies (CFPS) and the county-level gazetteer data we hand-collected from hundreds of books. The CFPS 2010 wave surveyed 33,600 respondents.[14] The survey covers a rich array of individual- and household-level information, including individual self-assessed anthropometric measurements. We focus on the height of individuals born between 1950 and 1990 and this restriction reduces our sample to a total of 25,243 individuals. Excluding individuals with missing values relevant to our estimation, reduces our final sample to 18,508 individuals. While our sampling frame is based on the 2010 wave, to reduce reporting errors on height, we take the mean height reported in the three survey years (2010, 2012 and 2014). But, to maintain the sample size we include individuals who reported height at least once in the three

[13] Here, we must stress that while OCP would mainly work through the fertility effect, economic reform and fertility may have some interactive impact on height. For example, the rural Household Responsibility System originally divided land use right according to number of household members, which is related to the household fertility. Later, this rule was gradually eroded in many regions.

[14] CFPS is a nationally representative household survey. As the survey was sampled at the household level, it may not be representative at the individual level. We compare the age and gender distributions of our sample (those who were born between 1950 and 1990) with the same age group of the 2010 Population Census data. We find that while the CFPS has 47.53% male population, the same ratio from the Census is 50.56%. The average age of the 20-60 year old population is 41.7 for males and 41.3 for females in the CFPS data, and 38.1 for males and 37.9 for females in the census data. The CFPS covers 162 counties/districts out of total of 2872 county/district level administrations.

waves. We also exclude obvious outliers from our final sample.[15]

The height variable in the CFPS is self-reported. Comparing self-reported height by birth cohort from the CFPS with the measured heights for the same cohorts from the China Health and Nutrition Survey (CHNS) we find that the differences between self-assessed and measured heights are small. However, the height-by-age measure could be subject to other biases. One is that there could be positive selection among those that survive to older ages, which could create upward bias among older cohorts. In the case of China, there is also a concern about those born just before and during the famine years.[16] For cohorts born after the early 1960s it should not be a problem. Another measurement-related issue is that height decreases in older age, largely due to compression of the spinal column, which would lead to a downward bias in earlier cohorts.[17] Both of these potential measurement issues point to our 1950s cohorts as being problematic, but the direction of the bias from the two issues are opposite and should to some extent cancel out. To allow analysis for the longest period we opt to keep the 1950s cohorts in the sample, but test the sensitivity of this inclusion for our main conclusions.[18]

An important feature of the CFPS is that it records each respondent's birthplace and household registration status in childhood at the county level, which the CHNS does not.[19] This means that we can accurately link individuals to local socioeconomic conditions during early childhood. We divide our final sample into two subsamples: those whose household registration at age 3 was rural (16,051) and those who were registered as urban (2,457). Given the vastly different economic institutions in rural and urban regions discussed above, we conduct separate analyses for rural and urban samples.

To link individuals to local socioeconomic conditions in early childhood, we merge the CFPS data with data collected from local gazetteers using birth year and birth county. Our sample individuals were born into 159 counties of 25 provinces.[20] The data from county gazetteers include annual agricultural and industrial output values (and total output values), government social expenditure (mainly including expenditure on health and education), mortality rates, and total population over

[15]Table A1 in Appendix A reports details as to how many observations were excluded based on each of the sample restrictions.

[16]Gørgens et al. (2012) estimate that selection during the Great Chinese Famine increased height by about 1-2cm. On the other hand, the so-called the scarring effect decreased height by about the same amount.

[17]Huang et al. (2013) estimate the height shrinkage for Chinese men and women from the age of 60 onwards at about 2 cm per decade. Further, Fernihough and McGovern (2015) found that the potential shrinkage at age 50 to 60 is quite small and stable.

[18]See Table A8 in Appendix A.

[19]At the time of the survey 2% of the rural sample and 15% of the urban sample was not living in their county of birth.

[20]Note that the number of counties discussed here is individuals' birth county, while in footnote 14 we presented the number of county surveyed in CFPS 2010. If use survey county definition, our sample came from 161 counties.

the years from 1950 to 1990. According to Almond et al. (2019), data from the gazetteers are similar to those from the yearbooks but are more accurate when they disagree. However, not all counties have all the variables for every year between 1950 and 1990. We use yearbook data to fill in some of the missing values, but that still leaves a rather large proportion of missing values. To ensure that the sample size is not reduced too much, we use available years of county-level data together with the prefecture or provincial time trend to obtain predictions for the years in which the county-level data are missing. We also test the sensitivity of our results to different imputation methods. More details of data sources, collection, and prediction methods are presented in Appendix C.

There is another issue related to merging county-level gazetteer data with individual-level data. In China each region has two types of household registration (*hukou*): 'rural' or 'urban'. Most of the population in rural areas have 'rural' household registrations, but a small proportion also has 'urban' household registrations, such as local mayors or a small number of public servants. In urban regions, most of the population has 'urban' household registrations, but there are also 'rural' *hukou* people living in the periphery of the cities. However, the aggregated data on county-level socioeconomic conditions do not distinguish between 'rural' and 'urban' within counties. Thus, we merge county-level variables with individuals' birth-county identification, regardless of their *hukou* status.[21] In the regressions, however, we include a variable which captures each county's rural or urban *hukou* population share. The information on rural or urban *hukou* populations at the county level is also from the county gazetteers.

4.2 Descriptive Statistics

Table 1 reports the summary statistics separately for the rural and urban samples. As height development is concentrated in early childhood, we define 'rural' and 'urban' samples based on individuals' *hukou* registration status at age 3. Panel A of Table 1 reports individual and parental characteristics, and Panel B presents the socioeconomic conditions of the individuals' birth counties in their birth year.

The average height of our rural male sample is 169 cm, and that of the urban sample is 172 cm, a 3 cm difference. For females, the difference is around 2 cm (=161-159). However, as indicated in Figure 1, the difference is larger for the younger birth cohorts than for the older birth cohorts. The average ages of the rural and urban samples are about the same, with the rural sample being around

[21] For example, individuals A and B were born in a county C. A's household registration status was urban, and B's was rural at birth. In our data, county-level characteristics for A are the same as for B.

0.7 years older than the urban sample. The average number of siblings is 3 among rural individuals and 2 for urban individuals.

One key element of family background is parental education. The average years of schooling of the parents (the average of mother and father) is 2.9 years for rural individuals and 6.2 years for urban individuals. To better capture family income level when our sample individuals were young, we generate a variable for father's occupation rank when the child was 14 years of age. We rank the 364 occupations based on the mean earnings in 2010 for each occupation.[22] The higher the father's occupation rank, the higher is average earnings; the average is 66 in the rural sample and 191 in the urban sample. 11.4% and 21.1% of individuals' fathers were communist party members in the rural and urban samples, respectively. The individual's own education differs sharply between rural and urban individuals by about 4.5 years of schooling. Finally, the proportion of Han Chinese is over 90% in both samples.

Panel B of Table 1 shows that people in rural areas faced poorer socioeconomic conditions compared with those in urban areas.[23] For example, the total value of output per capita in the rural sample is 650 yuan in 1990 prices, compared with 3,690 yuan in the urban sample.[24] Local government social expenditure per capita is 10 yuan for rural people, and 20 yuan for urban people. The average mortality rate is higher in the rural sample than in the urban sample.

5 Results

Before presenting our results it is worth stressing the following data limitations: (1) around 10-48% of county level information (depending on variables in question) are missing and we used imputed values for these missing values;[25] (2) it would be best to use household income at the critical ages, 0 to 3, to capture the household level economic conditions, but our data only allow us to use household income at age 14. These data limitations should be borne in mind when assessing the results presented below.

[22] It would be more accurate to use the average income level of each occupation at child's age 14, rather than as of 2010 but we do not have that information.

[23] Earlier, we mentioned that county-level variables are aggregated regardless of *hukou* status. Here we see significant differences in county-level characteristics because rural *hukou* individuals are more likely to live in rural areas where the socioeconomic environments are poorer than in urban areas.

[24] The reason why the urban sample also recorded agricultural output is that there are rural areas within each prefecture, such as the peripheries of cities. When Statistical Bureaus of these cities report their total output, it includes both non-agriculture and agriculture outputs.

[25] We conduct some sensitivity test with regard to the method used to predict for missing values in Table A4, Table A5 and Table A6.

Table 1: Descriptive statistics

	Rural *hukou*			Urban *hukou*		
Mean (Std. Dev)	Total	Male	Female	Total	Male	Female
Panel A: Individual characteristics						
Height (cm)	163.85	168.79	158.83	166.69	172.20	160.79
	(7.37)	(5.73)	(5.13)	(7.82)	(5.63)	(5.05)
Birth year	1969	1969	1969	1970	1970	1969
	(11.05)	(11.04)	(11.05)	(11.81)	(11.74)	(11.89)
Age in 2010	41.16	41.31	41.00	40.48	40.27	40.70
	(11.05)	(11.04)	(11.06)	(11.80)	(11.73)	(11.89)
Male	0.50			0.52		
	(0.50)			(0.50)		
Number of siblings	3.00	2.92	3.07	2.01	1.96	2.06
	(1.85)	(1.86)	(1.84)	(1.86)	(1.83)	(1.89)
Birth order	2.56	2.56	2.57	2.15	2.16	2.14
	(1.59)	(1.59)	(1.60)	(1.48)	(1.47)	(1.49)
Birth order index	0.06	0.09	0.03	0.15	0.18	0.11
	(1.21)	(1.20)	(1.22)	(0.95)	(0.95)	(0.95)
Parental average schooling years	2.94	2.95	2.92	6.20	6.24	6.17
	(3.23)	(3.23)	(3.23)	(4.29)	(4.32)	(4.25)
Father's occupation rank at child's age 14	66.14	65.40	66.89	191.23	190.06	192.49
	(80.27)	(79.49)	(81.04)	(79.52)	(79.40)	(79.66)
Father's party membership at child's age 14	0.11	0.11	0.11	0.21	0.21	0.22
	(0.32)	(0.32)	(0.32)	(0.41)	(0.40)	(0.41)
Individual schooling years	6.83	7.70	5.95	11.40	11.45	11.33
	(4.37)	(3.98)	(4.58)	(3.36)	(3.25)	(3.48)
Han Chinese	0.90	0.90	0.91	0.96	0.96	0.96
	(0.29)	(0.29)	(0.29)	(0.19)	(0.19)	(0.20)
Panel B: County characteristics[1)]						
Total output per capita[2)]	0.07			0.37		
	(0.10)			(0.39)		
Agricultural output per capita[2)]	0.03			0.03		
	(0.02)			(0.07)		
Industrial output per capita[2)]	0.03			0.33		
	(0.09)			(0.38)		
Gov. expenditure on social services per capita[2)]	0.001			0.002		
	(0.002)			(0.003)		
Death rate (‰)	9.52			7.27		
	(5.63)			(4.45)		
N	16051	8098	7953	2457	1270	1187

Note:
1) County-level variables are weighted by the rural *hukou* population share for the rural sample and urban *hukou* population share for the urban sample;
2) Unit: 10000 yuan; measured in 1990 prices.

5.1 Long-term trends in the Chinese heights

This section explores the long-term trend in adult heights of individuals born between 1950 and 1990. As shown in Figure 1, over the four decades, male height increased by around 6-7 cm, and female height increased by around 4 cm. A common pattern across the four groups is that upward trends are steeper among post-1970 birth cohorts relative to pre-1970. As discussed in Section 3 we estimate two different specifications of equation (1), where time is measured as: (1) a linear trend and (2) three decade dummy variables.

Panels A and B of Table 2 present the results for the rural and urban samples respectively. Columns 1 and 3 present results for the linear trend in height for males and females separately. The results in Panel A indicate that the annual height increase is 0.127 cm for rural males, and 0.070 cm for rural females. Panel B shows that annual height growth is 0.153cm for urban males and 0.104 for urban females, which is approximately 0.03 cm more than for the rural sample. Since birth-county fixed effects are included, this reflects within-county growth patterns. Thus, over 40 years, and controlling for regional variations, rural and urban male height increased by 5.08 and 6.12 cm, respectively, while rural and urban female height increased by 2.80 and 4.16 cm, respectively.

Columns 2 and 4 present the results using three dummy variables to capture the effect of changes in the economic and social environment over different decades on height. The omitted category in each case is those who were born in the 1950s. Thus, as each decade of the 1960s to 1980s is compared to the same base, their relative magnitudes should not be affected by the shrinkage issue due to the 1950s cohort being in their 50s. Panel A shows that rural males born in the 1960s are 0.944 cm taller than the reference category of the 1950s, while the difference for rural females is 1.122 cm. For the urban sample (Panel B) we find that 1960s-born males and females are 0.589 and 0.721 cm taller than their 1950s counterparts, respectively, but the difference is statistically significant only for the female sample. The fact that female 1960s cohorts are doing better than their male counterparts when compared to those born in the 1950s could be related to the shrinkage issue, which is greater for men than for women (see also Huang et al., 2013; Fernihough and McGovern, 2015).

The 1970s cohorts are taller than the 1960s cohorts for both rural and urban samples. The difference for rural males is 0.993 cm (=1.937-0.944), for rural females it is 0.247 cm (=1.369-1.122), for urban males it is 1.267 cm (=1.856-0.589), and for urban females it is 0.727 cm (=1.448-0.721). The differences between the 1970s and 1980s cohorts are 2.010 cm for rural males, 0.935 cm for rural females, 2.522 cm for urban males, and 1.898 cm for urban females. All of these differences

Table 2: Time trend in height growth

	Rural Total		Rural Male		Rural Female	
	(1)	(2)	(3)	(4)	(5)	(6)
Time trend	0.091***		0.127***		0.070***	
	(0.006)		(0.006)		(0.006)	
Born in 1960s		0.785***		0.944***		1.122***
		(0.138)		(0.153)		(0.157)
Born in 1970s		1.416***		1.937***		1.369***
		(0.177)		(0.199)		(0.167)
Born in 1980s		2.859***		3.947***		2.304***
		(0.202)		(0.203)		(0.188)
Constant	162.140***	162.672***	166.401***	167.218***	157.498***	157.657***
	(0.120)	(0.112)	(0.118)	(0.118)	(0.111)	(0.109)
Observations	16051	16051	8098	8098	7953	7953
Adjusted R^2	0.097	0.096	0.225	0.222	0.164	0.163

	Urban Total		Urban Male		Urban Female	
	(1)	(2)	(3)	(4)	(5)	(6)
Time trend	0.149***		0.154***		0.107***	
	(0.011)		(0.012)		(0.009)	
Born in 1960s		0.634		0.589		0.721**
		(0.382)		(0.422)		(0.359)
Born in 1970s		2.326***		1.856***		1.448***
		(0.497)		(0.526)		(0.540)
Born in 1980s		4.293***		4.378***		3.346***
		(0.350)		(0.408)		(0.361)
Constant	163.775***	164.892***	169.151***	170.463***	158.710***	159.422***
	(0.212)	(0.243)	(0.233)	(0.296)	(0.176)	(0.249)
Observations	2457	2457	1270	1270	1187	1187
Adjusted R^2	0.090	0.087	0.223	0.214	0.118	0.117

Note:
1) Robust Standard Errors (SEs) are presented in parentheses;
2) SEs are clustered at the county and prefecture levels for the rural and urban samples, respectively;
3) County and prefecture fixed effects are included for the rural and urban samples, respectively;
4) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

are statistically significant with the exception of urban females between the 1960s and 1970s.[26]

The results obtained from Table 2 confirm the strong and sustained increase in average Chinese heights, which exceeded 1 cm per decade for three out of four subgroups. Despite the economic and social turmoil that blighted much of the period, this is comparable with the experience of western Europe in the century to 1980 although it is slower than that of South Korea and Taiwan since the 1950s (Hatton, 2014; Schwekendiek and Baten, 2019). Most of the height growth for the Chinese population occurred after the 1970s and particularly so in the 1980s, which corresponds with the economic reforms from the late 1970s.

Also evident from Table 2 is that the urban population, on average, is taller than the rural population (the constant terms in Table 2 indicate that urban males and females who were born in the 1950s are about 3 cm and 2 cm taller than their counterparts born in rural areas, respectively[27]) and also had more height growth than their rural counterparts. China's rural-urban divide policy can probably explain a large part of the difference in the level and increment of average height between rural and urban populations. This is not only because urban average per capita income is almost three times that of the rural average and urban people consume more calories but also because healthcare provision has always been more widespread and better quality in urban than in rural areas as discussed in Section 2.

The final observation from Table 2 is that the average height of males increased by more than their female counterparts over time. More importantly, the gender gap in height is larger in rural than urban areas and increasingly so. This finding is consistent with Schwekendiek and Baten (2019) who also show that a pro-male bias in height is larger among the 1980s birth cohort than the 1970s cohort. The increasing difference in height between men and women over time, especially for the rural population, is likely to be related to the introduction of the OCP under the culture of son-preference. We investigate this issue further in Section 5.2.2.

[26] The significance test results are available upon request from the authors.

[27] When comparing the constant terms across regressions one needs to be aware of the omitted categories of multiple category dummy variables are consistent across the regressions. Luckily, the only group of dummy variables we include in these regression (apart from dummies for decades) is the county-fixed effects. We have estimated these regressions without county-fixed effects and the results show that they do not affect the constant terms much. These results are available upon request from the authors.

5.2 Height and its contributing factors

5.2.1 Socioeconomic factors

Although columns 2 and 4 of Table 2 have shown that average height growth is the strongest for those born during the late 1970s and the 1980s, what exactly contributed to this is unknown. In this subsection we turn to our second question to examine how changes in local living conditions are associated with the height-growth pattern. As previously discussed, equation (2) examines the effect of socioeconomic conditions in the individual's birth county in the first three years of life on adult height.

Table 3 presents the results for the rural total, male and female samples (columns 1 to 3) and the urban total, male and female samples (columns 4 to 6). Controlling for parental education, occupation, and party membership status when the person was a child (aged 14 years), individual and family characteristics, together with birth-county fixed effects, we find that local conditions during the first three years of life play important roles in rural individuals' adult height. For rural areas, every 10% increase in agricultural and industrial outputs increases average height by 0.043 cm and 0.021 cm, respectively, and the coefficients are statistically significant at the 5% and 1% levels, respectively. These are rather large impacts. During this period, China's average agricultural and industrial output increased by 3.1% and 8.0% per annum, respectively (NBS, 1999). In particular, during the economic reform period of the 1980s, annual growth rates were 6.0% and 7.6%, respectively. These large effects, however, mainly affected rural male average height, but not that of females. A 10% increase in agricultural and industrial outputs increased average male height by 0.075 cm and 0.026 cm and both are statistically significant, but the effects on average rural female height are almost zero and none is statistically significant.[28] Government social expenditure per capita (mainly on education and health) at the time of individuals' first three years of life also appear to have been an important contributor to both rural male and female average heights. However, the effect is almost twice the size for males as for females. A 10% increase in social spending increases male average height by 0.065cm and female average height by 0.033cm. General health conditions at the time of individuals' first three years of life, measured by the county-level mortality rate, has the expected negative effect overall but is significant only for females. A 10% reduction in the mortality rate is associated with a 0.0053 cm increase in rural female average height.

[28]The joint significance of the two variables is also not statistically significant with a p-value of 0.26. Table A2 in Appendix A includes total output per capita instead of the two separate output variables and shows a consistent pattern.

Table 3: Effects of regional social and economic factors on height

	Rural sample			Urban sample		
	Total	Male	Female	Total	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Log Agricultural output p.c.- age 0 to 3	0.427**	0.746***	0.074	0.039	0.212*	-0.155
	(0.168)	(0.233)	(0.217)	(0.106)	(0.125)	(0.137)
Log Industrial output p.c.- age 0 to 3	0.206***	0.259**	0.158	0.066	-0.057	0.177*
	(0.077)	(0.101)	(0.103)	(0.089)	(0.167)	(0.106)
Log Gov. Social expenditure p.c.- age 0 to 3	0.477***	0.646***	0.332***	0.893***	1.134***	0.542***
	(0.099)	(0.136)	(0.115)	(0.092)	(0.123)	(0.147)
Death rate (‰) - age 0 to 3	-0.032**	-0.015	-0.053***	0.019	0.018	0.014
	(0.012)	(0.015)	(0.016)	(0.037)	(0.047)	(0.052)
Parental schooling years	0.077***	0.102***	0.049**	0.127***	0.169***	0.123***
	(0.016)	(0.023)	(0.022)	(0.033)	(0.039)	(0.042)
Father's occupation rank at child's age 14	0.002***	0.002***	0.002**	0.000	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Father's party membership at child's age 14	0.360***	0.299	0.382**	-0.093	-0.329	0.148
	(0.129)	(0.195)	(0.190)	(0.294)	(0.409)	(0.353)
Male	10.022***			11.340***		
	(0.096)			(0.245)		
Han Chinese	0.108	0.259	-0.005	-0.459	0.145	-1.149
	(0.250)	(0.365)	(0.329)	(0.848)	(0.842)	(1.020)
Number of siblings	-0.054*	-0.054	-0.028	-0.226***	-0.137	-0.315***
	(0.028)	(0.036)	(0.039)	(0.069)	(0.098)	(0.097)
Birth order index	0.061	0.034	0.053	-0.058	-0.064	-0.042
	(0.037)	(0.050)	(0.050)	(0.121)	(0.219)	(0.177)
Observations	16051	8098	7953	2457	1270	1187
Adjusted R^2	0.560	0.225	0.164	0.610	0.218	0.119

Note:

1) Robust Standard Errors (SEs) are presented in parentheses;

2) SEs are clustered at the county and prefecture levels for the rural and urban samples;

3) Unit: log of 10000 yuan; measured in 1990 price level

4) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;

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

These results suggest that changes in the local socioeconomic environment contributed more to the heights of rural males than to rural females. Indeed, including the four county-level variables in rural male and female regressions explains, respectively, 23% and 16% of the average height of rural males and females (the adjusted R^2s are 0.23 and 0.16). If we exclude the four social/economic environment variables, the adjusted R^2 for the rural male sample is reduced by 2.5 percentage points to 0.20 (or by 11%), whereas for the rural female sample, it only drops by 1 percentage point to 0.155 (6.1% of the total explained variation).

In contrast, for the urban sample, the output variables have limited effects on average height, except for a small positive effect of industrial output on the average heights of urban females. The effect of government social spending per capita, however, seems to be a very important influence on heights of urban people. A 10% increase in government social spending increases average urban height by 0.089 cm. The effect is particularly strong for males; a 10% increase in social spending increases average urban male height by 0.113 cm, whereas for females, it increases average height by only half as much: 0.054 cm. The level of government social spending per capita for our sample prefectures increased by 8.4% per annum between 1950 and 1990. For the last decade, the annual increase was much higher at 13.9% per year. On the other hand, the reduction in mortality rates plays no role in average urban height, either for males or females. Once again, the four local socioeconomic condition variables, in particular government social spending, contributed a rather large proportion of explained height variation for the male sample (11.8%=0.0257/0.218) but much less for the female sample (5.3%=0.006/0.119).

Parental characteristics at the time the person was aged 14 are used to capture family socioeconomic background when the individual was young. We have three variables: the average years of schooling of the mother and father (labelled as parental years of schooling); father's occupational rank when the child was aged 14, based on the mean earnings of the occupation in 2010; and whether the father was a communist party member when the child was 14. All three variables are statistically significant and positively related to individuals' height for the rural sample. For the urban sample, the only statistically significant parental variable is parental years of schooling. Other family and individual characteristics are not associated with average height apart from the number of siblings. For both rural and urban full samples, sibship size is associated with individual's height negatively, which is consistent with the Quantity-Quality trade-off theory, as well as the empirical literature on other countries (Hatton and Martin, 2010; Liu, 2014; Hatton, 2014, 2017; Bailey et al., 2016; Kubo and Chaudhuri, 2017). However, when we separately estimate for the male and female samples, none of the coefficients is statistically significant. As discussed previously, while the fertility policy in China has changed significantly, Table 3 does not take the policy change into account.

5.2.2 Economic reform, the One Child Policy and height

In light of the dramatic shifts in public policy outlined above, it seems likely that the effects of some of the variables that appear in Table 3 would have changed over time, especially during the economic reform era. As the economic reform took place at the same time as the OCP, it is difficult to separate out the two effects. However, we can make inferences by first deriving a measure of exposure to the economic reform, controlling for the famine and CR events, and then interacting the reform exposure with sibship to distinguish the effects of the OCP from those of economic reforms as represented by equation 3.

To recap, the exposure variables capture the share of the individual's first three years of life exposed to each of the three events. For exposure to the economic reform it measures the shift from the planned economy to the market system. The introduction of the Household Responsibility System in rural areas and the subsequent switching from a planned to a market-oriented economic system for the urban economy led China to experience unprecedented and sustained economic growth; the abolition of the commune system reduced social services previously provided in rural areas; and the introduction of the OCP substantially reduced fertility. These policy changes would have had important implications for average height. An increase in income provides families with more resources for everyone, while at the same time the reduction in social services at the village level and the privatisation of healthcare in cities could have the opposite effect. Similarly, the introduction of the OCP reduced fertility and gave each child access to more resources, *ceteris paribus*. However, under the culture of son-preference, and with the financial punishment for a violation of the policy being very high,[29] the net impact of the OCP is complicated depending on whether the individual is male or female, and whether the family has son-preference.

The lack of social-welfare provision resulting from the abolition of the old commune system during the initial economic reform period mostly affected rural people. The potential negative impact of OCP was also mainly borne by rural people, in particular, rural girls. In general, rural Chinese are more traditional and have stronger son-preference than urban dwellers. In the presence of strong son preference, and when the firstborn is a girl, it is more likely that parents would have a second birth. And if the second is still a girl, they may continue to have a third and more births until they have a boy, whereas if the first or the second birth is a boy, parents are likely to stop. Thus, girls are likely to have more siblings than boys under the OCP. In addition, girls are more likely to be born in families that violate the birth quota and hence trigger a substantial fine, which in turn generates more financial difficulties for the family. Therefore, the introduction of the economic reform and the OCP might have contributed to the lower height growth for rural people, girls in particular, than their urban counterparts.

[29] According to Scharping (2003) and Ebenstein (2010), the fine for violation of OCP in the 1980s was between 20% and 360% of urban household annual income, which is normally 2 to 3 times rural household annual income.

Table 4: Reform and OCP effects — Rural sample

	Rural Total			Rural Male			Rural Female		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log Agricultural output p.c.- age 0 to 3	0.427**	0.194	0.193	0.746***	0.401	0.399	0.074	-0.050	-0.050
	(0.168)	(0.172)	(0.170)	(0.233)	(0.247)	(0.244)	(0.217)	(0.235)	(0.235)
Log Industrial output p.c.- age 0 to 3	0.206***	0.180**	0.182**	0.259**	0.230**	0.226**	0.158	0.141	0.144
	(0.077)	(0.079)	(0.078)	(0.101)	(0.099)	(0.098)	(0.103)	(0.105)	(0.105)
Log Gov. Social expenditure p.c.- age 0 to 3	0.477***	0.358***	0.349***	0.646***	0.472***	0.466***	0.332***	0.267**	0.261**
	(0.099)	(0.099)	(0.099)	(0.136)	(0.137)	(0.137)	(0.115)	(0.122)	(0.123)
Death rate (‰) - age 0 to 3	-0.032**	-0.032**	-0.036***	-0.015	-0.014	-0.019	-0.053***	-0.053***	-0.055***
	(0.012)	(0.013)	(0.013)	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)
Parental schooling years	0.077***	0.071***	0.067***	0.102***	0.095***	0.090***	0.049**	0.046**	0.043**
	(0.016)	(0.016)	(0.016)	(0.023)	(0.023)	(0.024)	(0.022)	(0.022)	(0.022)
Father's occupation rank at child's age 14	0.002***	0.002***	0.002***	0.002***	0.002***	0.002**	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father's party membership at child's age 14	0.360***	0.386***	0.400***	0.299	0.334*	0.357*	0.382**	0.397**	0.401**
	(0.129)	(0.130)	(0.131)	(0.195)	(0.196)	(0.197)	(0.190)	(0.191)	(0.190)
Male	10.022***	10.024***	10.006***						
	(0.096)	(0.096)	(0.095)						
Han Chinese	0.108	0.110	0.097	0.259	0.246	0.244	-0.005	0.002	-0.012
	(0.250)	(0.252)	(0.255)	(0.365)	(0.366)	(0.363)	(0.329)	(0.332)	(0.335)
Number of siblings	-0.054*	-0.040	0.002	-0.054	-0.031	0.012	-0.028	-0.021	0.002
	(0.028)	(0.028)	(0.029)	(0.036)	(0.037)	(0.037)	(0.039)	(0.040)	(0.041)
Birth order index	0.061	0.072*	0.072*	0.034	0.049	0.054	0.053	0.059	0.058
	(0.037)	(0.037)	(0.037)	(0.050)	(0.050)	(0.049)	(0.050)	(0.051)	(0.050)
Exposure age 0 to 3		0.738***	1.455***		1.066***	1.812***		0.405*	0.814**
		(0.164)	(0.218)		(0.248)	(0.322)		(0.244)	(0.319)
Exposure age 0 to 3 × Number of siblings			-0.393***			-0.440***			-0.210**
			(0.068)			(0.108)			(0.095)
Observations	16051	16051	16051	8098	8098	8098	7953	7953	7953
Adjusted R^2	0.560	0.560	0.561	0.225	0.227	0.228	0.164	0.164	0.164

Note:
1) County fixed effects, a share of rural population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;
2) Robust Std. Err. are clustered at the county level;
3) * (**,***) denotes statistical significance at the 10% (5%, 1%) level;
4) P values of joint significance test for 'Number of siblings' and 'Exposure age 0 to 3 × Number of siblings' are 0.000 in column 3 and 6 and 0.082 in column 9;

Table 5: Reform and OCP effects — Urban sample

	Urban Total			Urban Male			Urban Female		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log Agricultural output p.c.- age 0 to 3	0.039	-0.130*	-0.130*	0.212*	0.040	0.043	-0.155	-0.341***	-0.345***
	(0.106)	(0.073)	(0.072)	(0.125)	(0.093)	(0.093)	(0.137)	(0.108)	(0.108)
Log Industrial output p.c.- age 0 to 3	0.066	0.029	0.031	-0.057	-0.094	-0.093	0.177*	0.142	0.144
	(0.089)	(0.099)	(0.096)	(0.167)	(0.179)	(0.174)	(0.106)	(0.096)	(0.096)
Log Gov. Social expenditure p.c.- age 0 to 3	0.893***	0.463***	0.455***	1.134***	0.687***	0.684***	0.542***	0.091	0.079
	(0.092)	(0.121)	(0.122)	(0.123)	(0.199)	(0.195)	(0.147)	(0.184)	(0.187)
Death rate (‰) - age 0 to 3	0.019	0.021	0.014	0.018	0.020	0.014	0.014	0.014	0.009
	(0.037)	(0.033)	(0.033)	(0.047)	(0.048)	(0.048)	(0.052)	(0.047)	(0.046)
Parental schooling years	0.127***	0.114***	0.112***	0.169***	0.156***	0.152***	0.123***	0.107***	0.106***
	(0.033)	(0.031)	(0.031)	(0.039)	(0.039)	(0.039)	(0.042)	(0.039)	(0.039)
Father's occupation rank at child's age 14	0.000	0.001	0.001	-0.001	-0.000	-0.000	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Father's party membership at child's age 14	-0.093	0.002	0.003	-0.329	-0.280	-0.280	0.148	0.314	0.312
	(0.294)	(0.307)	(0.307)	(0.409)	(0.408)	(0.408)	(0.353)	(0.378)	(0.379)
Male	11.340***	11.324***	11.322***						
	(0.245)	(0.238)	(0.238)						
Han Chinese	-0.459	-0.384	-0.405	0.145	0.087	0.049	-1.149	-0.949	-0.947
	(0.848)	(0.827)	(0.822)	(0.842)	(0.834)	(0.823)	(1.020)	(1.003)	(1.005)
Number of siblings	-0.226***	-0.167**	-0.143**	-0.137	-0.072	-0.051	-0.315***	-0.260***	-0.240**
	(0.069)	(0.067)	(0.068)	(0.098)	(0.099)	(0.103)	(0.097)	(0.096)	(0.092)
Birth order index	-0.058	-0.030	-0.031	-0.064	-0.029	-0.021	-0.042	-0.014	-0.022
	(0.121)	(0.118)	(0.117)	(0.219)	(0.211)	(0.211)	(0.177)	(0.171)	(0.171)
Exposure age 0 to 3		1.731***	1.934***		1.834***	2.016***		1.787***	1.963***
		(0.308)	(0.329)		(0.579)	(0.589)		(0.584)	(0.628)
Exposure age 0 to 3 × Number of siblings			-0.391*			-0.390			-0.316
			(0.230)			(0.278)			(0.399)
Observations	2457	2457	2457	1270	1270	1270	1187	1187	1187
Adjusted R^2	0.610	0.613	0.613	0.218	0.224	0.224	0.119	0.125	0.125

Note:
1) County fixed effects, a share of rural population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;
2) Robust Std. Err. are clustered at the county level;
3) * (**,***) denotes statistical significance at the 10% (5%, 1%) level;
4) P values of joint significance test for 'Number of siblings' and 'Exposure age 0 to 3 × Number of siblings' are 0.019 in column 3, 0.211 in column 6, and 0.030 in column 9;

To understand the economic reform and the OCP effects, equation (3) is estimated for the rural and urban samples separately and the results for the two samples are reported in Tables 4 and 5, respectively. The first three columns of each table report the results for the total sample, while columns 4 to 6, and 7 to 9 present results for male and female samples separately. For all three samples we present three specifications: the reproduction of the results from equation (2) (column 1) for comparison purposes; a specification adding only $Expo_{ic}$, the post-1979 exposure variable, to equation (2) to capture the net effect of both economic reform and the introduction of the OCP; and the estimated results from equation (3).

In this subsection we discuss the results from the second specification. The results for both the rural and urban samples show that people who spent some of their first 3 years of life after the introduction of economic reform and OCP are on average taller than those who did not (compare columns 2, 5 and 8 with 1, 4 and 7). For the rural male sample, this difference is around 1.07cm. That is to say, those who spent 100% of their first 3 years of life during the reform and OCP era are 1.07cm taller than those whose first three years of life was during the pre-reform and pre-OCP era. This difference for the rural female sample is smaller, at around 0.41cm, but is still statistically significant. The difference for the urban male sample is the largest, at 1.83cm. It is more than twice as large (1.79cm) for the urban female sample as for the rural female sample. As a proportion of mean heights the gains are 0.6% for rural males 0.3% for rural females, 1.1% for urban males, and 1.1% for urban females. Thus, the proportional gains are greater for men than for women, at least in rural areas, and greater for the urban-born than for the rural-born.

Table 6: Selected results including interaction terms between $Expo_{ic}$ and county variables

	Rural sample	Urban sample
	(1)	(2)
(a) Log Agricultural output p.c.- age 0 to 3	0.110	-0.153**
	(0.186)	(0.076)
(b) Log Industrial output p.c.- age 0 to 3	0.181**	0.024
	(0.080)	(0.104)
(c) Log Gov. social expenditure p.c.- age 0 to 3	0.361***	0.279*
	(0.108)	(0.160)
(d) Death rate (‰) - age 0 to 3	-0.035***	-0.004
	(0.013)	(0.032)
Exposure age 0 to 3	2.031*	5.130***
	(1.136)	(1.813)
(e) Exposure age 0 to 3 × Log Agricultural output p.c.- age 0 to 3	0.118	-0.105
	(0.211)	(0.105)
(f) Exposure age 0 to 3 × Log Industrial output p.c.- age 0 to 3	0.047	0.144
	(0.078)	(0.140)
(g) Exposure age 0 to 3 × Log Gov. social expenditure p.c.- age 0 to 3	-0.024	0.571***
	(0.147)	(0.211)
(h) Exposure age 0 to 3 × Death rate (‰) - age 0 to 3	-0.141***	0.012
	(0.038)	(0.124)
Observations	16051	2457
Adjusted R^2	0.561	0.614
Total effect for the exposed group		
Coefficient (a)+(e)	0.228	-0.258
Test h_0 : Coefficient on $(a) + (e) = 0$ [P value]	[0.294]	[0.028]
Coefficient (b)+(f)	0.228	0.168
Test h_0 : Coefficient on $(b) + (f) = 0$ [P value]	[0.036]	[0.137]
Coefficient on (c)+(g)	0.337	0.850
Test h_0 : Coefficient on $(c) + (g) = 0$ [P value]	[0.018]	[0.000]
Coefficient on (d)+(h)	-0.176	0.008
Test h_0 : Coefficient on $(d) + (h) = 0$ [P value]	[0.000]	[0.950]

Note:

1) Robust Standard Errors (SEs) are presented in parentheses;

2) SEs are clustered at the county and prefecture levels for the rural and urban samples;

3) Unit: log of 10000 yuan; measured in 1990 price level;

4) All the covariates included in Table 3 are included;

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

It is interesting to note that when the exposure variable is included, the coefficients on all of the previous statistically significant production output variables become smaller in size and some even become statistically insignificant. The magnitude of the coefficients on 'per capita government social spending' are also reduced in size by around one third relative to the estimation without controlling for the exposure variable. These results suggest that perhaps the gains in height due to local socioeconomic conditions are more important after the economic reform and the introduction of OCP. Indeed, when we interact the exposure variable with the local socioeconomic condition variables, we observe, for the rural sample, that the slope for industrial output is steeper for the group whose first 3 years of life was during the reform era than for those who were born earlier (see Table 6).[30] For the earlier cohort, this coefficient is positive and statistically significant. The effect of agricultural output is also positive for both groups and it is larger for the younger group, but none are precisely estimated. The effect of per capita government spending on health and education is strong and positive for the older group but slightly smaller for the younger group. Both are statistically significant. The largest difference is in the negative effect of mortality on height. While it is significant for both groups, the magnitude is much larger for the post-reform group. For the urban sample, however, the only statistically significant difference between the post- and pre-reform groups is the effect of government per capita health and education spending. For pre- and post-reform groups, the coefficient on government health and education spending is positive, but it is statistically significantly larger for the post-reform group. It is also very interesting that when we interact these socioeconomic variables with the exposure variable, the size of the exposure coefficient increases. This indicates that there are economic reform effects over and above the increase in output and reduction in mortality, perhaps due to the increased ability of households to allocate resources in their own interest.

5.3 The fertility effect

The discussion above does not try to separate the economic reform effect from that of the introduction of the OCP. Studies have shown that fertility decline improves child health through the Quantity-Quality trade-off mechanism (Becker and Lewis, 1973; Liu, 2014; Kubo and Chaudhuri, 2017). As family size decreases, the amount of food or other resources available for each child increases. This may lead to an increase in height. At the same time, the introduction of the OCP was accompanied by a significant financial fine to the households that violated their birth quotas. Because of this, it is likely that in the post-OCP era, the difference between children who grew up in households with fewer children and those with many children would be larger than in the pre-OCP era. In columns 3, 6 and 9 of Tables 4 and 5 we try to separately gauge the economic reform and the introduction of

[30] Table 6 reports selected results from estimating a modified version of equation 3, where we add the interaction terms between $Expo_{ic}$ and the full set of local economic conditions (E_{cn}) but not including the interaction between $Expo_{ic}$ and 'siblings'. The full results of these estimations are available upon request from the authors.

the OCP effects by including an interaction term between the exposure variable and the number-of-siblings variable. The interaction term between the exposure and sibling variables measures the net effect on the 1979-and-post cohorts of the reduced fertility due to the OCP and the negative effect due to potential fines incurred for violation of the OCP birth quotas.[31] The results in Table 4 for both rural male and female samples show that the coefficients on the number of siblings are small and not statistically significant. The interaction terms, however, are large and precisely estimated. These results suggest that height of rural male and female individuals who were born before the introduction of the OCP were not affected by the number of siblings they had, but if they were born after the OCP an additional sibling reduces male and female heights by 0.43cm and 0.21cm, respectively. The differences in the coefficients for the male and female samples suggest that a given reduction in sibship size due to the OCP benefited males more than females.[32]

Furthermore, after controlling for potential OCP effects (working through the number of siblings and its interaction with post-1979 exposure, $(Expo_{ic} * Sib_{ic})$), the coefficients on the main exposure variable, which should capture the economic reform effect, now become larger. The effect can be expressed as $\Delta H_{ic}/\Delta Expo_{ic} = \gamma_5 + \gamma_6 Sib_{ic}$ derived from equation (3). Intuitively, for rural males, who have no sibling, exposed to the economic reform since birth are 1.81cm taller than those who were not exposed the economic reform during the first 3 years of life (pre-reform birth cohorts). This effect falls as the sibling number increases. At the mean numbers of siblings for the post-1979 cohort,[33] 1.63 siblings, the height is 1.09cm (1.81+(-0.44*1.63)) taller than their counterparts born before 1979. Thus, for a single child, this is an increase of 70% over the effect estimated without controlling for the OCP effect, while for those with the mean number of siblings, it is 3.0% (1.093/1.066-1) higher.

For rural females, the economic reform effect is 0.81cm and 0.47cm for a single child, and for people with 1.63 siblings, respectively. Relative to the estimated results without controlling for the interaction term, the estimated results for a single child and those with mean level of siblings are 100% and 16% (0.47/0.41-1), respectively, higher than the estimates without controlling for the

[31] We also estimated another specification using a dummy variable for birth year at and after 1979 to interact with the number of siblings variable and the results are largely consistent with the results presented here. These results are available upon request from the authors.

[32] Males also benefited more because, as noted above, their number of siblings fell by more after the reform.

[33] At the median number of siblings for rural males of the post-1979 cohort (1 sibling), the effect is 1.37. Note that we evaluate the impact at the post-1979 cohort median and mean level, rather than the overall median or mean level of siblings. The reason for doing so is because the economic reform effect is experienced only by the post-reform cohorts. For this group, the overall median or mean number of siblings is not that meaningful. For example, the overall mean number of siblings for our rural sample as a whole is 3, whereas only 18% of the post-reform cohort has 3 or more siblings.

interaction term.

For the urban samples, the coefficients on the interactions between exposure and sibling numbers for both urban male and female samples are sizeable, but they are not precisely estimated and joint significance tests between the 'sibling' level effect and the interaction term for the male sample is not statistically significant either. Controlling for the OCP effect, the coefficient on exposure increases slightly from 1.86cm to 2.02cm for males, a 9% increase, and from 1.79cm to 1.96cm for females, a 9.5% increase. As the median number of siblings for the urban post-1979 cohort is zero, these large effects are for a sizeable group of urban people.

Our results from columns 3, 6 and 9 of Tables 4 and 5 suggest that the net economic reform effect on height is large in both rural and urban areas. But compared to the mean height, urban dwellers benefited much more than their rural counterparts and rural females gained the least (the size of the coefficient on the 'exposure' variable obtained in estimation of equation (3) as a share of the mean height is 1.2% for urban males and females, 1.1% for rural males, and 0.5% for rural females).

In addition, we also calculate the net effect on height of the reduction in sibling size for the cohort born after the introduction of the OCP relative to their counterparts born before the OCP. Table 7 presents the results for rural and urban population by gender groups. Column 1 exhibits the effect of the fertility on height ($\Delta H_{ic}/\Delta Sib_{ic}$), columns 2-4 show the change in the average number of siblings between the pre- and post-OCP cohorts, column 5 calculates the total effect on height due to reduction of sibling size, while column 6 indicates the change in height as proportion of the average height for each group due to the change in the number of siblings between the pre- and post-OCP cohorts. As can be seen from the table, the net effect of the reduction in sibling size is larger in urban areas than in rural areas and rural females gained the least from reduced sibling size.

Table 7: Calculated effect on height of reduction in sibling size between cohorts born pre- and post-OCP

	$\gamma_4 + \gamma_6$ (1)	pre-Sib (2)	post-Sib (3)	Δ in Sib (4)=(3)−(2)	Total effect (5)=(1)×(4)	% of mean height (6)
			Rural Sample			
Males	-0.428	3.32	1.49	-1.83	0.78	0.46%
Females	-0.208	3.45	1.76	-1.69	0.35	0.22%
			Urban Sample			
Males	-0.441	2.59	0.38	-2.21	0.98	0.57%
Females	-0.556	2.72	0.42	-2.30	1.28	0.80%

6 Robustness checks

6.1 Endogeneity of fertility

Until now we have assumed the 'number of siblings' is exogenous. However, it is likely that some unobservable parental characteristics determined both the individual's height and the number of siblings he/she has. For example, parental son-preference may affect how many children parents choose to have as well as how family resources are allocated among their male/female children. If so, our OLS coefficients on the number-of-siblings variable would be biased. To test whether the endogeneity of the number-of-siblings may alter our conclusion, we estimate equation (2) using an instrumental variable (IV) approach. The following first-stage equation is estimated:

$$Sib_{ict} = \gamma_0 + \gamma_1 \sum_{n=0}^{2} \omega_n E_{c(t_n)} + \gamma_2 X^P_{i(t_{14})} + \gamma_3 W_{ict} + \gamma_4 Z_{ct_m} + \rho_c + \upsilon_{ict} \tag{4}$$

where t_m indicates individual's birth year, and Z is the instrument. The instrument is a birth-control rate at the county level for each year since the 1950, constructed by Chae (2020), where birth-control methods include male and female sterilisation, intrauterine devices, pills, use of condoms, and any other method of birth control. Given that the birth-control rate increased sharply after the introduction of the OCP, if we take the Local Average Treatment Effect (LATE) interpretation of the IV results, the IV estimate of β_4 in equation (2) measures the average sibling effect generated from the increase in the birth-control rate, or generated by the introduction of the OCP. Appendix E provides details of how the birth-control rate is constructed.[34]

Our IV is highly correlated with the endogenous variable (the number of siblings) as indicated at the bottom of Table 8 (as well as Table A3 in Appendix A and Figure E2 in Appendix E). We argue that the share of fertile women using contraceptives at the county level should not have a direct impact on individuals' adult height over and above its impact through the number-of-siblings variable.

Table 8 reports the selected IV results. The coefficient on the number of siblings for different samples is still negative and statistically significant, supporting our conclusion using OLS. The magnitude of the coefficients, however, increases substantially. This may be related to the LATE interpretation: those whose parents' fertility decision was altered by the change in the birth control rate in the county of residence had a larger effect on their height than the average effect for individuals

[34] Also see Chapter 2 in Chae (2020) for further details

whose parents' fertility decision were not affected by the birth control policies.

Due to the fact that both our IV and exposure variables have a strong time trend we are unable to estimate the specification with the exposure effect included.

Table 8: Instrumental variable results for sibship size effect on height

	Rural sample			Urban sample		
	Total	Male	Female	Total	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Log agricultural output p.c.- age 0 to 3	-0.008	0.085	-0.121	-0.338**	-0.149	-0.522**
	(0.177)	(0.274)	(0.244)	(0.135)	(0.160)	(0.229)
Log industrial output p.c.- age 0 to 3	0.116	0.098	0.124	0.009	-0.132	0.151
	(0.084)	(0.105)	(0.108)	(0.124)	(0.235)	(0.099)
Log gov. social expenditure p.c.- age 0 to 3	0.388***	0.499***	0.293**	0.362**	0.596***	0.047
	(0.102)	(0.135)	(0.119)	(0.163)	(0.184)	(0.340)
Death rate (‰) - age 0 to 3	-0.017	0.009	-0.046***	0.063	0.070	0.048
	(0.013)	(0.016)	(0.017)	(0.047)	(0.046)	(0.064)
Parental schooling years	0.038**	0.040	0.032	0.040	0.067	0.057
	(0.018)	(0.028)	(0.025)	(0.039)	(0.045)	(0.045)
Father's occupation rank at child's age 14	0.002***	0.002***	0.002**	0.001	-0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Father's party membership at child's age 14	0.502***	0.513**	0.447**	-0.158	-0.399	0.059
	(0.136)	(0.200)	(0.194)	(0.299)	(0.379)	(0.342)
Male	9.928***			11.211***		
	(0.095)			(0.223)		
Han Chinese	0.146	0.301	0.013	-0.464	-0.129	-0.941
	(0.249)	(0.352)	(0.327)	(0.943)	(1.024)	(1.043)
Birth order index	0.162***	0.212***	0.094*	0.379**	0.368	0.364
	(0.042)	(0.059)	(0.057)	(0.182)	(0.252)	(0.279)
Number of siblings	-0.587***	-0.871***	-0.269*	-1.202***	-1.141***	-1.197**
	(0.117)	(0.180)	(0.158)	(0.249)	(0.288)	(0.471)
Observations	16051	8098	7953	2457	1270	1187
Adjusted R^2	0.547	0.175	0.158	0.582	0.164	0.058
Kleibergen-Paap Wald F stat	357.104	260.281	236.240	188.931	217.200	84.618

Note:

1) Robust Standard Errors (SEs) are presented in parentheses;

2) SEs are clustered at the county and prefecture levels for the rural and urban samples;

3) Unit: log of 10000 yuan; measured in 1990 price level

4) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;

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

6.2 Missing values in local gazetteer data

As discussed in the data section and detailed in Appendix C, our county-level data have many missing values. We used available years of county-level data on a particular variable together with the prefecture/provincial level trend for this variable to predict values for the missing years.

To check whether our main results are driven by these predicted values we estimate equation (2) excluding all the predicted observations for each county-level variable in turn for the rural sample. Table A4 and Table A5 in Appendix A present the estimated results for equation (2) and equation (3), respectively.[35] These results show that dropping in turn the predicted observations for agricultural and industrial output has little effect on the estimated coefficients for these variables, although their statistical significance declines slightly. The coefficients on the other variables are also little affected. When predicted values of death rates and government social expenditure variables are excluded, the magnitudes of the coefficients increase for both males and females.

Table A6 in Appendix A reproduces the Table 4 results excluding all the predicted values, i.e. individuals with at least one predicted value of county-level variables are excluded. This involves a dramatic 37% reduction in the number of observations. Nevertheless, the results remain broadly similar to those in Table 4. In particular, the coefficients on government social expenditure retain their size and significance. The coefficients on the exposure variable and its interaction with sibship size are similar in magnitude and significance to those in Table 4.

6.3 Other sensitivity tests

As mentioned earlier, parental height is an important determinant of child height, and our main equations (2) and (3) omit this variable due to the limited information available. We observe parental height information only when the individual's parents also participate in the CFPS survey. Out of 16,051 observations in the rural sample, only 3,141 individuals (19.6%) have mother's height and 2,664 (16.6%) have father's height information. We lose more than 90% of female and about 70% of male observations, and the huge gender difference is likely due to the fact that males are more likely than females to live with their parents in the same households after marriage.

In Table A7 we compare the results with and without the mother's height variable for the same subsample with parental height information.[36] The table shows that controlling for parental height does not greatly affect the results. The coefficients on government social expenditure become smaller once mother's height is controlled for (compare columns 5 and 6 with columns 2 and 3) but are still statistically significant at the 5% or 10% levels and still sizeable. Regardless of the inclusion of the mother's height variable, the coefficients on the number of siblings are statistically significant and

[35] Results for urban samples are available upon request from the authors.

[36] We choose mother's height as it has about 500 more observations. The estimation results with father's height variable are available upon request from the authors.

the magnitudes do not change much. In addition, adding mother's height increases the magnitude and significance level of the exposure effect and its interaction term with sibling variable.[37]

Another concern is the potential shrinkage from age 50 to 60 in our sample as discussed in the Data Section. Table A8 in Appendix A shows that excluding those over 50 reduces the coefficient size of agricultural output to a large extent (columns 1 and 4), although it is still economically meaningful among males (column 5). Among males, the coefficient on the government social expenditure increases by 69% (columns 2 and 5), and the sibling effect on height almost doubles when people over 50 are excluded.

Also, in our main estimation, we focused on the economic reform impact only as pre-economic reform Chinese economic structure did not change much, despite the political and economic upheavals experienced during the Great Famine and the Cultural Revolution period. Here we test if adding controls for these two events would change our results. Table A9 in Appendix A presents the comparison of the estimation of equation (3) with and without the famine and CR exposure variables.[38] The results indicate that, just as we expected earlier, neither of the additional events has a statistically significant negative effect on height. If anything, the CR exposure has a positive and statistically significant effect on height of rural women. More importantly including these two event exposure variables does not change our results.

Finally, in the main estimations we used local conditions during the individual's first three years to examine the correlation between covariates and their height. As the first three years height growth only accounts for 40% of individuals' adult height, one might argue that the window should be widened. To examine the sensitivity of this decision, we use ages 0-5 (accounting for 50% of total growth) and ages 0-7 (60% of total growth) as exposure weights for the relevant covariates. The results are reported in Tables A10 and A11 in Appendix A. They show that changing the adjustment factor does not seem to affect our main conclusions.

7 Conclusion

This study explores the relationship between the heights of individuals born between 1950 and 1990 and economic development and fertility policy changes over the period. We find that China experienced rapid growth in the average height of individuals. If we take a linear trend, rural and

[37]The results from estimating these regressions for males and females separately are available upon request.

[38]The exposure variables for the famine and CR are defined the same way as that for the exposure to the post-1978 variable.

urban male heights increased by more than 1 cm per decade (1.3 cm for rural males and 1.5 cm for urban males), which is faster than what was found for the Western European males between 1870 and 1970 (1 cm per decade). The height trend, however, was not constant over the 40-year period, with the fastest growth occurring in the 1980s. In the 1980s decade, the height of rural and urban males grew by 2 cm and 2.5 cm, respectively. For rural and urban females the 1980s decade saw 0.9 cm and 1.7 cm growth, respectively.

The sharp increase in height in the 1980s cohort reflects two forces: the introduction of the market-oriented economic reform and the introduction of the OCP. We find that the height growth during this period is mainly driven by the economic reforms, which led to fast and sustained improvements in local socioeconomic conditions. Controlling for the introduction of the OCP, we observe an even larger height growth due to the reforms (at the mean number of siblings, – 1.63 sibling for the rural sample; 0.4 sibling for the urban sample –, the increase is 1.09 cm for rural males, 0.47 cm for rural females and around 2.0 cm for both urban males and females). When separately examining the periods before and after 1979, we find that the effects of local economic conditions on individuals' height are stronger for the post-1979 era than for the pre-1979 era. For the pre-reform era, the single most important factor for height growth is per capita government spending on health and education.

This study provides suggestive evidence as to how the market-oriented economic reforms changed people's lives as indicated by the significant increase in the average height of the population. Economic reform improved the living standards of rural families by boosting agricultural and local industrial output growth, which should have directly increased food availability and household income. By expanding the choice set faced by households the reform also improved health and well-being over and above that due to the improvement in the available local resources. However, the rural people gained the least from the OCP, especially rural females, possibly due to the wide-spread son-preference in rural society.[39]

[39] As we only have county level data between 1950-90, using age 0-5 or 0-7 as the adjustment factors would significantly reduce our sample size. As a compromise we use provincial-level variables to predict for the missing value in county level socioeconomic condition variables. For each of the county level input variables, we first estimate the equation $e_{c,p,t} = \alpha_0 + \alpha_1 e_{p,t} + \delta_t + u_{c,p,t}$ using data for 1950-1990, where $e_{c,p,t}$ is a county-level socioeconomic condition variable of county c in province p at year t. We then use the estimated $\hat{\alpha}_0$ and $\hat{\alpha}_1$ to predict missing values for the period from 1991 to 1996 by calculating $\hat{e}_{c,p,t} = \alpha_0 + \alpha_1 e_{p,t}$.

References

Almond, D., Li, H., and Zhang, S. (2019). Land reform and sex selection in China. *Journal of Political Economy*, 127(2):560–585.

Ashton, B., Hill, K., Piazza, A., and Zeitz, R. (1984). Famine in China, 1958-61. *Population and Development Review*, 10(4):613–645.

Bailey, R. E., Hatton, T. J., and Inwood, K. (2016). Health, height, and the household at the turn of the twentieth century. *Economic History Review*, 69(1):35–53.

Becker, G. S. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81(2):S279–S288.

Bettelheim, C. (1988). Economic reform in China. *Journal of Development Studies*, 24(4):15–49.

Bi, Z.-W. and Ji, C.-Y. (2005). Secular growth changes in body height and weight in children and adolescents in Shandong, China between 1939 and 2000. *Annals of Human Biology*, 32(5):650–665.

Bloom, G. and Xingyuan, G. (1997). Health sector reform: Lessons from China. *Social Science & Medicine*, 45(3):351–360.

Blumenthal, D. and Hsiao, W. (2005). Privatization and its discontents - the evolving Chinese health care system. *New England Journal of Medicine*, 353(11):1165–1170.

Booth, A., Meng, X., Fan, E., and Zhang, D. (2018). The intergenerational behavioural consequences of a socio-political upheaval. *IZA Discussion Paper*, 11991.

Borensztein, E. and Ostry, J. D. (1996). Accounting for China's growth performance. *American Economic Review, Papers and Proceedings*, 86(2):224–228.

Cai, Y. (2010). China's below-replacement fertility: Government policy or socioeconomic development? *Population and Development Review*, 36(3):419–440.

Center for Population Studies, C. A. o. S. S. and of the China's Population Yearbook, E. B. (1986). The major events of China's population activities. In *The China's Population Yearbook (1985)*, pages 1263–1288. Chinese Social Sciences Press, Beijing.

Chae, M. (2020). *Essays on Childhood Environments and Human Capital Formation*. PhD thesis, The Australian National University.

Chen, Y. and Zhou, L.-A. (2007). The long-term health and economic consequences of the 1959-1961 famine in China. *Journal of Health Economics*, 26:659–681.

Chow, G. C. (1993). Capital formation and economic growth in China. *Quarterly Journal of Economics*, 108(3):809–842.

Cole, T. J. and Wright, C. M. (2011). A chart to predict adult height from a child's current height. *Annals of Human Biology*, 38:662–668.

Dong, Z. and Phillips, R. M. (2008). Evolution of China's health-care system. *The Lancet*, 372:1715–1716.

Doyle, O., Harmon, C. P., Heckman, J. J., and Tremblay, R. E. (2009). Investing in early human development: Timing and economic efficiency. *Economics and Human Biology*, 7(1):1–6.

Du, S. F., Wang, H. J., Zhai, F. Y., and Popkin, B. M. (2014). China in the period of transition from scarcity and extensive undernutrition to emerging nutrition-related non-communicable diseases, 1949-1992. *Obesity Review*, 15(Suppl 1):8–15.

Ebenstein, A. (2010). The "missing girls" of China and the unintended consequences of the one child policy. *Journal of Human Resources*, 45(1):87–115.

Edlund, L., Li, H., Yi, J., and Zhang, J. (2013). Sex ratios and crime: Evidence from China's one-child policy. *Review of Economics and Statistics*, 95(5):1520–1534.

Feeney, G. and Wang, F. (1993). Parity progression and birth intervals in China: The influence of policy in hastening fertility decline. *Population and Development Review*, 19(1):61–101.

Fernihough, A. and McGovern, M. E. (2015). Physical stature decline and the health status of the elderly population in England. *Economics and Human Biology*, 16:30–44.

Gietel-Basten, S., Han, X., and Cheng, Y. (2019). Assessing the impact of the "one-child policy" in China: A synthetic control approach. *PLOS ONE*, 14(11).

Gong, Y.-L. and Chao, L.-M. (1982). The role of barefoot doctors. *American Journal of Public Health*, 72(9 Supplement):59–61.

Gørgens, T., Meng, X., and Vaithianathan, R. (2012). Stunting and selection effects of famine: A case study of the Great Chinese Famine. *Journal of Development Economics*, 97:99–111.

Hatton, T. J. (2014). How have europeans grown so tall? *Oxford Economic Papers*, 66:349–372.

Hatton, T. J. (2017). Stature and sibship: Historical evidence. *History of the Family*, 22(2-3):175–195.

Hatton, T. J. and Martin, R. M. (2010). The effects on stature of poverty, family size, and birth order: British children in the 1930s. *Oxford Economic Papers*, 62:157–184.

Hesketh, T. and Wei, X. Z. (1997). Health in China: From Mao to market reform. *BMJ*, 24(314):1543.

Hipgrave, D. (2011). Communicable disease control in China: From Mao to now. *Journal of Global Health*, 1(2):224–238.

Huang, W., Lei, X., Ridd, G., Strauss, J., and Zhao, Y. (2013). Health, height, height shrinkage, and SES at older ages: Evidence from China. *American Economic Journal: Applied Economics*, 5(2):86–121.

Huang, Y. (2011). The sick man of Asia - China's health crisis. *Foreign Affairs*, 90(6):119–136.

Kubo, M. and Chaudhuri, A. (2017). Gender gap in health status of children in the context of one-child policy in China: Is it sibling rivalry or son preference? *Journal of Family and Economic Issues*, 38:204–217.

Li, W. and Yang, D. T. (2005). The Great Leap Forward: Anatomy of a central planning disaster. *Journal of Political Economy*, 113(4):840–877.

Liang, M. H., Eichling, P. S., Fine, L. J., and Annas, G. J. (1973). Chinese health care: Determinants of the system. *Am J Public Health*, 63(2):102–110.

Liang, Y. and Gibson, J. (2018). Do siblings take your food away? using china's one-child policy to test for child quantity-quality trade-offs. *China Economic Review*, 48:14–26.

Lin, J. Y. (1988). The household responsibility system in China's rural reform. In *Proceedings of the Twentieth International Conference of Agricultural Economists*. International Association of Agricultural Economists, Buenos Aires.

Lin, J. Y. (1990). Collectivization and China's agricultural crisis in 1959-1961. *Journal of Political Economy*, 98(6):1228–1252.

Lin, J. Y. (1992). Rural reforms and agricultural growth in China. *American Economic Review*, 82(1):34–51.

Liu, H. (2014). The quality-quantity trade-off: evidence from the relaxation of China's one-child policy. *Journal of Population Economics*, 27:565–602.

Liu, X. and Cao, H. (1992). China's cooperative medical system: Its historical transformations and the trend of development. *Journal of Public Health Policy*, 13(4):501–511.

Liu, Y. (2004). Development of the rural health insurance system in China. *Health Policy and Planning*, 19(3):159–165.

Liu, Y., Hsiao, W. C., Li, Q., Liu, X., and Ren, M. (1995). Transformation of China's rural health care financing. *Social Science & Medicine*, 41(8):1085–1093.

MacFarquhar, R. and Schoenhals, M. (2006). *Mao's Last Revolution*. Harvard University Press.

Mamidi, R. S., Kulkarni, B., and Singh, A. (2011). Secular trends in height in different states of india in relation to socioeconomic characteristics and dietary intakes. *Food and Nutrition Bulletin*, 32:23–34.

McCormick, M., Brooks-Gunn, J., Buka, S., Goldman, J., Yu, J., Salganik, M., Scott, D., Bennett, F., Kay, L., Bernbaum, J., and Bauer, C. (2006). Early intervention in low birth weight premature infants: results at 18 years of age for the infant health and development program. *Pediatrics*, 117(3):771–780.

Meng, X. (2000). *Labour market reform in China*. Cambridge University Press.

Meng, X. (2012). Labor market outcomes and reforms in China. *Journal of Economic Perspectives*, 26(4):75–102.

Meng, X., Qian, N., and Yared, P. (2015). The institutional causes of China's great famine, 1959–1961. *Review of Economic Studies*, 82:1568–1611.

Morgan, S. L. (2000). Richer and taller: Stature and living standards in China, 1979-1995. *The China Journal*, 44:1–39.

Morgan, S. L. (2004). Economic growth and the biological standard of living in china, 1880–1930. *Economics and Human Biology*, 2:197–218.

NBS (1999). Comprehensive statistical data and materials on 50 years of new China, National Bureau of Statistics.

NBS (2010). China compendium of statistisc 1949-2010.

NCD (2020). Height and body-mass index trajectories of school-aged children and adolescents from 1985 to 2019 in 200 countries and territories: a pooled analysis of 2181 population-based studies with 65 million participants. *The Lancet*, 396:1151–1524.

Onis, M. d., Garza, C., Victora, C. G., Onyango, A. W., Frongillo, E. A., and Martines, J. (2004). The who multicentre growth reference study: Planning, study design, and methodology. *Food and Nutrition Bulletin*, 25(1):S3–S84.

Peng, X. (1991). *Demographic Transition in China*. Oxford: Clarendon Press.

Perkins, J. M., Khan, K. T., Smith, D., G., and Subramanian, S. V. (2011). Patterns and trends of adult height in india in 2005-2006. *Economics and Human Biology*, 9:184–193.

Perkins, J. M., Subramanian, S. V., Davey Smith, G., and Özaltin, E. (2016). Adult height, nutrition, and population health. *Nutrition Reviews*, 74(3):149–165.

Qian, N. (2009). Quantity-quality and the one child policy: The only-child disadvantage in school enrollment in rural china. *NBER Working Paper 14973*.

Ravallion, M. and Chen, S. (2007). China's (uneven) progress against poverty. *Journal of Development Economics*, 1(82):1–42.

Salaff, J. W. (1973). Mortality decline in the People's Republic of China and the United States. *Population Studies*, 27(3):551–576.

Scharping, T. (2003). *Birth control in China 1949-2000: population policy and demographic development*. Routledge, London.

Schwekendiek, D. and Baten, J. (2019). Height development of men and women from China, South Korea, and Taiwan during the rapid economic transformation period of the 1960s-1980s. *Economics and Human Biology*, 34:169–180.

Steckel, R. H. (2009). Heights and human welfare: Recent developments and new directions. *Explorations in Economic History*, 46(1):1–23.

Subramanian, S. V., Özaltin, E., and Finlay, J. E. (2011). Height of nations: A socioeconomic analysis of cohort differences and patterns among women in 54 low- to middle-income countries. *PLoS ONE*, 6(4):1–13.

Tu, J. (2016). The lost generation - “barefoot doctors” in post-reform China. *China Perspectives*, 4:7–17.

Wang, F. (2011). The future of a demographic overachiever: Long-term implications of the demographic transition in China. *Population and Development Review*, 37(Supplement):173–190.

Wang, X. (2018). China’s macroeconomics in the 40 years of reform. In Ross Garnaut, L. S. and Fang, C., editors, *China’s 40 Years of Reform and Development 1978-2018*. ANU Press.

Weng, Q., Gao, X., He, H., Li, S. (2019) Family size, birth order and educational attainment: Evidence from China *China Economic Review*, 57 (201346).

Xin, H.-C., Amsbary, J., and Powell, L. (2014). Urban health insurance reform in China. *International Conference on Economic Management and Trade Cooperation (EMTC 2014)*.

Yip, W. and Hsiao, W. C. (2015). What drove the cycles of Chinese health system reforms? *Health Systems & Reform*, 1(1):52–16.

Zhang, D. and Unschuld, P. U. (2008). China’s barefoot doctor: past, present, and future. *The Lancet*, 372(9653):1865–1867.

Zhang, J., Xu, P., and Li, F. (2020). One-child policy and childhood obesity. *China Economic Review*, 59:100938.

Zhu, N., Ling, Z., Lane, J., and Hu, S. (1989). Factors associated with the decline of the cooperative medical system and barefoot doctors in rural China. *Bulletin of the World Health Organization*, 67(4):431–441.

Appendix A: Figures and Tables

Table A1: Main sample generation

Description	Observations
All CFPS respondents in 2010	33600
+ Born between 1950 and 1990	25243
++ Height info. available	25095
+++ *hukou* at age 3 info. available	24793
++++ Birth county info. available	19430
+++++ Birth county gazetteer data available	18715
++++++ Individual gender and ethnicity info. available	18677
+++++++ Sibling info. available	18511
++++++++ Height>190cm excluded	18508
Final sample - Rural	16051
Final sample - Urban	2457

Table A2: Estimated results with total output

	Rural sample			Urban sample		
	Total	Male	Female	Total	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Log total output p.c. - age 0 to 3	0.520***	0.812***	0.228	0.185	0.046	0.309*
	(0.145)	(0.197)	(0.188)	(0.131)	(0.156)	(0.183)
Log Gov. social expenditure p.c.- age 0 to 3	0.546***	0.724***	0.386***	0.874***	1.099***	0.532***
	(0.097)	(0.128)	(0.117)	(0.087)	(0.117)	(0.162)
Death rate (‰) - age 0 to 3	-0.035***	-0.018	-0.054***	0.021	0.019	0.019
	(0.013)	(0.016)	(0.016)	(0.037)	(0.046)	(0.051)
Parental schooling years	0.079***	0.106***	0.050**	0.125***	0.172***	0.114***
	(0.016)	(0.024)	(0.022)	(0.033)	(0.038)	(0.043)
Father's occupation rank at child's age 14	0.002***	0.002**	0.002**	0.000	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Father's party membership at child's age 14	0.354***	0.284	0.382**	-0.091	-0.361	0.202
	(0.130)	(0.195)	(0.190)	(0.288)	(0.408)	(0.346)
Male	10.023***			11.339***		
	(0.096)			(0.242)		
Han Chinese	0.115	0.287	-0.009	-0.449	0.083	-1.121
	(0.249)	(0.360)	(0.329)	(0.847)	(0.844)	(1.017)
Number of siblings	-0.056*	-0.057	-0.028	-0.223***	-0.159	-0.283***
	(0.028)	(0.035)	(0.040)	(0.066)	(0.097)	(0.092)
Birth order index	0.064*	0.038	0.056	-0.061	-0.049	-0.051
	(0.037)	(0.050)	(0.050)	(0.122)	(0.222)	(0.177)
Observations	16051	8098	7953	2457	1270	1187
Adjusted R^2	0.560	0.224	0.164	0.610	0.218	0.120

Note:

1) Robust Standard Errors (SEs) are presented in parentheses;

2) SEs are clustered at the county and prefecture levels for the rural and urban samples;

3) Unit: log of 10000 yuan; measured in 1990 price level

4) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;

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

Table A3: IV first stage results

	Rural sample			Urban sample		
	Total	Male	Female	Total	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Birth control rate (%)	-0.021***	-0.021***	-0.022***	-0.022***	-0.022***	-0.023***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Observations	16051	8098	7953	2457	1270	1187
Kleibergen-Paap Wald F stat	357.104	260.281	236.240	188.931	217.200	84.618

Note:

1) Robust Standard Errors (SEs) are presented in parentheses;

2) SEs are clustered at the county and prefecture levels for the rural and urban samples;

3) Unit: log of 10000 yuan; measured in 1990 price level

4) All the variables in Table 8 are included;

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

Table A4: Main estimation excluding predicted observations (excluding predicted observations of one variable at a time)

	Main sample			Excluding predicted observations of Agricultural output			Industrial output			Death rate			Government social expenditure		
	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Log Agricultural output p.c.- age 0 to 3	0.427**	0.746***	0.074	0.393*	0.768**	0.042	0.422*	0.722**	0.118	0.329*	0.763***	-0.152	0.367*	0.737***	-0.012
	(0.168)	(0.233)	(0.217)	(0.224)	(0.316)	(0.307)	(0.242)	(0.315)	(0.310)	(0.188)	(0.272)	(0.225)	(0.192)	(0.275)	(0.250)
Log Industrial output p.c.- age 0 to 3	0.206***	0.259**	0.158	0.209*	0.331**	0.105	0.220**	0.250*	0.204	0.180**	0.198*	0.170	0.062	0.186	-0.045
	(0.077)	(0.101)	(0.103)	(0.110)	(0.154)	(0.147)	(0.099)	(0.134)	(0.132)	(0.086)	(0.115)	(0.117)	(0.097)	(0.138)	(0.121)
Log Gov. social expenditure p.c.- age 0 to 3	0.477***	0.646***	0.332***	0.481***	0.706***	0.249	0.449***	0.724***	0.186	0.593***	0.823***	0.409***	0.673***	0.915***	0.472***
	(0.099)	(0.136)	(0.115)	(0.152)	(0.203)	(0.174)	(0.156)	(0.213)	(0.170)	(0.129)	(0.169)	(0.142)	(0.117)	(0.150)	(0.147)
Death rate (‰) - age 0 to 3	-0.032**	-0.015	-0.053***	-0.008	0.006	-0.028	-0.012	0.004	-0.033*	-0.042***	-0.026	-0.061***	-0.015	0.013	-0.047***
	(0.012)	(0.015)	(0.016)	(0.015)	(0.019)	(0.021)	(0.017)	(0.022)	(0.019)	(0.014)	(0.017)	(0.018)	(0.013)	(0.016)	(0.016)
Parental schooling years	0.077***	0.102***	0.049**	0.081***	0.087***	0.075**	0.085***	0.120***	0.044	0.064***	0.074***	0.051**	0.085***	0.095***	0.069***
	(0.016)	(0.023)	(0.022)	(0.022)	(0.030)	(0.030)	(0.021)	(0.031)	(0.028)	(0.018)	(0.025)	(0.024)	(0.018)	(0.027)	(0.025)
Father's occupation rank at child's age 14	0.002***	0.002***	0.002**	0.003***	0.003***	0.002*	0.001*	0.002*	0.001	0.002***	0.002***	0.002**	0.002***	0.002**	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father's party membership at child's age 14	0.360***	0.299	0.382**	0.357*	0.280	0.384	0.423**	0.328	0.463*	0.371***	0.371*	0.326*	0.451***	0.411*	0.433*
	(0.129)	(0.195)	(0.190)	(0.189)	(0.303)	(0.284)	(0.165)	(0.280)	(0.246)	(0.135)	(0.221)	(0.196)	(0.154)	(0.241)	(0.228)
Male	10.022***			10.164***			10.044***			9.987***			9.994***		
	(0.096)			(0.126)			(0.126)			(0.098)			(0.110)		
Han Chinese	0.108	0.259	-0.005	0.091	-0.200	0.483	0.070	-0.094	0.298	0.005	0.043	0.019	0.309	0.339	0.366
	(0.250)	(0.365)	(0.329)	(0.289)	(0.500)	(0.457)	(0.354)	(0.488)	(0.420)	(0.280)	(0.396)	(0.365)	(0.275)	(0.359)	(0.383)
Number of siblings	-0.054*	-0.054	-0.028	-0.049	-0.026	-0.052	-0.055	-0.037	-0.052	-0.059*	-0.050	-0.045	-0.046	-0.026	-0.040
	(0.028)	(0.036)	(0.039)	(0.036)	(0.044)	(0.056)	(0.036)	(0.042)	(0.054)	(0.032)	(0.041)	(0.041)	(0.033)	(0.039)	(0.049)
Birth order index	0.061	0.034	0.053	0.025	0.015	0.016	0.072	0.069	0.047	0.056	0.017	0.069	0.060	0.049	0.036
	(0.037)	(0.050)	(0.050)	(0.051)	(0.078)	(0.066)	(0.045)	(0.068)	(0.061)	(0.040)	(0.055)	(0.053)	(0.043)	(0.057)	(0.058)
Observations	16051	8098	7953	8480	4227	4253	9209	4650	4559	13462	6781	6681	12043	6087	5956
Adjusted R^2	0.560	0.225	0.164	0.567	0.213	0.158	0.562	0.238	0.168	0.559	0.218	0.164	0.560	0.237	0.169

Note:

1) Rural sample only

2) Robust Standard Errors (SEs) are presented in parentheses;

3) SEs are clustered at the county and prefecture levels for the rural and urban samples;

4) Unit: log of 10000 yuan; measured in 1990 price level

5) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;

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

Table A5: Main estimation excluding predicted observations (excluding predicted observations of one variable at a time)

	Main sample			Excluding predicted observations of: Agricultural output			Industrial output			Death rate			Government social expenditure		
	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Log Agricultural output p.c.- age 0 to 3	0.193	0.399	-0.050	0.128	0.363	-0.108	0.186	0.385	-0.030	0.162	0.518*	-0.231	0.213	0.526*	-0.123
	(0.170)	(0.244)	(0.235)	(0.222)	(0.323)	(0.326)	(0.244)	(0.327)	(0.331)	(0.187)	(0.279)	(0.240)	(0.195)	(0.283)	(0.274)
Log Industrial output p.c.- age 0 to 3	0.182**	0.226**	0.144	0.192*	0.296**	0.099	0.214**	0.241*	0.201	0.170*	0.185*	0.167	0.036	0.149	-0.063
	(0.078)	(0.098)	(0.105)	(0.111)	(0.148)	(0.150)	(0.099)	(0.130)	(0.134)	(0.087)	(0.111)	(0.118)	(0.096)	(0.135)	(0.121)
Log Gov. Social expenditure p.c.- age 0 to 3	0.349***	0.466***	0.261**	0.330**	0.497**	0.159	0.299*	0.513**	0.092	0.448***	0.621***	0.334**	0.547***	0.746***	0.380**
	(0.099)	(0.137)	(0.123)	(0.151)	(0.205)	(0.185)	(0.157)	(0.222)	(0.183)	(0.132)	(0.177)	(0.154)	(0.127)	(0.161)	(0.159)
Death rate (‰) - age 0 to 3	-0.036***	-0.019	-0.055***	-0.013	0.002	-0.032	-0.018	-0.000	-0.039**	-0.046***	-0.030*	-0.064***	-0.020	0.007	-0.049***
	(0.013)	(0.016)	(0.017)	(0.015)	(0.019)	(0.021)	(0.017)	(0.023)	(0.019)	(0.015)	(0.017)	(0.018)	(0.013)	(0.016)	(0.017)
Parental schooling years	0.067***	0.090***	0.043**	0.069***	0.073**	0.067**	0.074***	0.109***	0.035	0.057***	0.067***	0.047**	0.077***	0.086***	0.064**
	(0.016)	(0.024)	(0.022)	(0.021)	(0.030)	(0.030)	(0.021)	(0.031)	(0.028)	(0.018)	(0.025)	(0.023)	(0.018)	(0.027)	(0.025)
Father's occupation rank at child's age 14	0.002***	0.002**	0.002**	0.002***	0.003***	0.002*	0.001*	0.002*	0.001	0.002***	0.002**	0.002**	0.002***	0.002*	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father's party membership at child's age 14	0.400***	0.357*	0.401**	0.413**	0.369	0.411	0.468***	0.402	0.483*	0.401***	0.417*	0.338*	0.490***	0.467*	0.456**
	(0.131)	(0.197)	(0.190)	(0.194)	(0.307)	(0.286)	(0.168)	(0.282)	(0.246)	(0.137)	(0.223)	(0.196)	(0.156)	(0.244)	(0.227)
Male	10.006***			10.138***			10.018***			9.973***			9.978***		
	(0.095)			(0.125)			(0.124)			(0.097)			(0.109)		
Han Chinese	0.097	0.244	-0.012	0.082	-0.209	0.468	0.049	-0.129	0.284	0.005	0.039	0.018	0.300	0.318	0.369
	(0.255)	(0.363)	(0.335)	(0.286)	(0.492)	(0.459)	(0.360)	(0.488)	(0.431)	(0.284)	(0.393)	(0.369)	(0.279)	(0.360)	(0.389)
Number of siblings	0.002	0.012	0.002	0.014	0.043	-0.012	0.005	0.028	-0.006	-0.013	-0.003	-0.014	0.004	0.032	-0.012
	(0.029)	(0.037)	(0.041)	(0.038)	(0.044)	(0.061)	(0.037)	(0.040)	(0.058)	(0.032)	(0.042)	(0.043)	(0.033)	(0.039)	(0.052)
Exposure age 0 to 3	1.455***	1.812***	0.814**	1.585***	1.939***	0.980**	1.537***	1.781***	1.123***	1.237***	1.429***	0.770**	1.298***	1.549***	0.822**
	(0.218)	(0.322)	(0.319)	(0.265)	(0.398)	(0.434)	(0.259)	(0.390)	(0.418)	(0.240)	(0.351)	(0.335)	(0.266)	(0.372)	(0.381)
Exposure age 0 to 3 × Number of siblings	-0.393***	-0.440***	-0.210**	-0.420***	-0.405**	-0.266**	-0.451***	-0.444***	-0.350***	-0.364***	-0.366***	-0.247**	-0.384***	-0.444***	-0.200*
	(0.068)	(0.108)	(0.095)	(0.097)	(0.159)	(0.130)	(0.085)	(0.137)	(0.129)	(0.082)	(0.121)	(0.108)	(0.078)	(0.119)	(0.107)
Birth order index	0.072*	0.054	0.058	0.033	0.034	0.017	0.078*	0.084	0.047	0.064	0.033	0.071	0.067	0.062	0.041
	(0.037)	(0.049)	(0.050)	(0.051)	(0.077)	(0.067)	(0.044)	(0.066)	(0.063)	(0.041)	(0.055)	(0.054)	(0.043)	(0.057)	(0.059)
Observations	16051	8098	7953	8480	4227	4253	9209	4650	4559	13462	6781	6681	12043	6087	5956
Adjusted R^2	0.561	0.228	0.164	0.568	0.217	0.159	0.564	0.241	0.169	0.560	0.220	0.165	0.561	0.239	0.169

Note:

1) Rural sample only

2) Robust Standard Errors (SEs) are presented in parentheses;

3) SEs are clustered at the county and prefecture levels for the rural and urban samples;

4) Unit: log of 10000 yuan; measured in 1990 price level

5) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;

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

Table A6: Main estimation excluding predicted observations (excluding all predicted values)

	Main sample			All predicted values are excluded								
	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log Agricultural output p.c.- age 0 to 3	0.427**	0.746***	0.074	0.046	0.344	-0.310	-0.063	0.154	-0.340	-0.048	0.178	-0.336
	(0.168)	(0.233)	(0.217)	(0.273)	(0.413)	(0.342)	(0.274)	(0.421)	(0.369)	(0.276)	(0.424)	(0.369)
Log Industrial output p.c.- age 0 to 3	0.206***	0.259**	0.158	0.148	0.325*	0.012	0.132	0.297*	0.008	0.143	0.294*	0.023
	(0.077)	(0.101)	(0.103)	(0.113)	(0.173)	(0.149)	(0.116)	(0.173)	(0.149)	(0.116)	(0.170)	(0.151)
Log Gov. Social expenditure p.c.- age 0 to 3	0.477***	0.646***	0.332***	0.881***	1.088***	0.674***	0.741***	0.828***	0.638**	0.731***	0.814***	0.628**
	(0.099)	(0.136)	(0.115)	(0.208)	(0.270)	(0.239)	(0.208)	(0.265)	(0.265)	(0.207)	(0.262)	(0.264)
Death rate (‰) - age 0 to 3	-0.032**	-0.015	-0.053***	-0.010	0.014	-0.039	-0.010	0.013	-0.040	-0.016	0.007	-0.045
	(0.012)	(0.015)	(0.016)	(0.022)	(0.026)	(0.032)	(0.022)	(0.026)	(0.032)	(0.023)	(0.026)	(0.032)
Parental schooling years	0.077***	0.102***	0.049**	0.067***	0.066*	0.074**	0.064**	0.063*	0.073**	0.059**	0.058	0.067*
	(0.016)	(0.023)	(0.022)	(0.025)	(0.037)	(0.035)	(0.025)	(0.037)	(0.035)	(0.025)	(0.036)	(0.034)
Father's occupation rank at child's age 14	0.002***	0.002***	0.002**	0.003***	0.003**	0.002	0.003***	0.003**	0.002	0.002***	0.003**	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father's party membership at child's age 14	0.360***	0.299	0.382**	0.587***	0.451	0.629**	0.608***	0.494	0.634**	0.624***	0.529	0.637**
	(0.129)	(0.195)	(0.190)	(0.205)	(0.378)	(0.308)	(0.205)	(0.375)	(0.304)	(0.206)	(0.378)	(0.302)
Male	10.022***			10.289***			10.286***			10.259***		
	(0.096)			(0.145)			(0.146)			(0.146)		
Han Chinese	0.108	0.259	-0.005	-0.097	-0.383	0.195	-0.084	-0.390	0.203	-0.094	-0.387	0.185
	(0.250)	(0.365)	(0.329)	(0.317)	(0.507)	(0.459)	(0.317)	(0.508)	(0.462)	(0.319)	(0.505)	(0.466)
Number of siblings	-0.054*	-0.054	-0.028	-0.055	-0.050	-0.035	-0.048	-0.034	-0.033	0.005	0.017	0.009
	(0.028)	(0.036)	(0.039)	(0.050)	(0.060)	(0.078)	(0.050)	(0.059)	(0.077)	(0.050)	(0.056)	(0.084)
Birth order index	0.061	0.034	0.053	0.079	0.059	0.090	0.085	0.067	0.091	0.078	0.069	0.082
	(0.037)	(0.050)	(0.050)	(0.067)	(0.094)	(0.084)	(0.067)	(0.095)	(0.084)	(0.067)	(0.094)	(0.086)
Exposure age 0 to 3							0.503**	0.919**	0.133	1.300***	1.782***	0.753
							(0.253)	(0.439)	(0.375)	(0.298)	(0.516)	(0.500)
Exposure age 0 to 3 × Number of siblings										-0.483***	-0.564***	-0.348**
										(0.114)	(0.189)	(0.147)
Observations	16051	8098	7953	5879	2923	2956	5879	2923	2956	5879	2923	2956
Adjusted R^2	0.560	0.225	0.164	0.583	0.231	0.176	0.583	0.232	0.176	0.584	0.234	0.177

Note:

1) Rural sample only

2) Robust Standard Errors (SEs) are presented in parentheses;

3) SEs are clustered at the county and prefecture levels for the rural and urban samples;

4) Unit: log of 10000 yuan; measured in 1990 price level

5) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;

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

Table A7: Estimated results with parental height variable included — Rural sample

	without parental height			with parental height		
	(1)	(2)	(3)	(4)	(5)	(6)
Log Agricultural output p.c.- age 0 to 3	0.950**	0.907**	0.935**	0.876**	0.864**	0.907**
	(0.414)	(0.426)	(0.424)	(0.383)	(0.393)	(0.384)
Log Industrial output p.c.- age 0 to 3	0.042	0.037	0.039	-0.050	-0.052	-0.051
	(0.212)	(0.213)	(0.211)	(0.210)	(0.212)	(0.208)
Log Gov. Social expenditure p.c.- age 0 to 3	0.802***	0.727**	0.709**	0.578**	0.556*	0.528*
	(0.281)	(0.295)	(0.298)	(0.271)	(0.282)	(0.284)
Death rate (‰) - age 0 to 3	-0.010	-0.008	-0.019	-0.008	-0.007	-0.023
	(0.041)	(0.041)	(0.041)	(0.039)	(0.039)	(0.039)
Parental schooling years	0.003	0.003	0.000	-0.024	-0.024	-0.029
	(0.034)	(0.034)	(0.035)	(0.033)	(0.033)	(0.033)
Father's occupation rank at child's age 14	0.002*	0.002*	0.002*	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father's party membership at child's age 14	0.619**	0.619**	0.623**	0.548*	0.548*	0.554*
	(0.296)	(0.296)	(0.295)	(0.290)	(0.290)	(0.288)
Male	11.025***	11.020***	11.002***	11.046***	11.045***	11.018***
	(0.266)	(0.266)	(0.266)	(0.263)	(0.261)	(0.260)
Han Chinese	0.286	0.281	0.268	0.385	0.384	0.365
	(0.501)	(0.501)	(0.496)	(0.462)	(0.463)	(0.451)
Number of siblings	-0.209***	-0.204***	-0.103	-0.156*	-0.155*	0.001
	(0.074)	(0.074)	(0.088)	(0.079)	(0.079)	(0.090)
Birth order index	0.079	0.082	0.074	0.149	0.150	0.139
	(0.102)	(0.103)	(0.102)	(0.106)	(0.108)	(0.106)
Exposure age 0 to 3		0.215	0.698		0.062	0.798*
		(0.385)	(0.489)		(0.364)	(0.453)
Number of siblings × Exposure age 0 to 3			-0.267*			-0.407***
			(0.149)			(0.139)
Mother's height (cm)				0.296***	0.296***	0.299***
				(0.024)	(0.024)	(0.024)
Observations	3141	3141	3141	3141	3141	3141
Adjusted R^2	0.507	0.507	0.508	0.553	0.553	0.554

Note:

1) Rural sample only

2) Robust Standard Errors (SEs) are presented in parentheses;

3) SEs are clustered at the county and prefecture levels for the rural and urban samples;

4) Unit: log of 10000 yuan; measured in 1990 price level

5) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;

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

Table A8: Main estimation excluding 1950s cohort — Rural sample

	Including 1950s cohort			Excluding 1950s cohort			Including 1950s cohort			Excluding 1950s cohort		
	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log Agricultural output p.c.- age 0 to 3	0.427**	0.746***	0.074	0.217	0.279	0.071	0.193	0.399	-0.050	0.120	0.191	-0.032
	(0.168)	(0.233)	(0.217)	(0.182)	(0.268)	(0.284)	(0.170)	(0.244)	(0.235)	(0.182)	(0.269)	(0.288)
Log Industrial output p.c.- age 0 to 3	0.206***	0.259**	0.158	0.148	0.260**	0.089	0.182**	0.226**	0.144	0.149	0.258**	0.088
	(0.077)	(0.101)	(0.103)	(0.092)	(0.105)	(0.149)	(0.078)	(0.098)	(0.105)	(0.094)	(0.103)	(0.153)
Log Gov. Social expenditure p.c.- age 0 to 3	0.477***	0.646***	0.332***	0.664***	1.079***	0.296*	0.349***	0.466***	0.261**	0.344**	0.784***	-0.025
	(0.099)	(0.136)	(0.115)	(0.123)	(0.192)	(0.160)	(0.099)	(0.137)	(0.123)	(0.145)	(0.256)	(0.201)
Death rate (‰) - age 0 to 3	-0.032**	-0.015	-0.053***	-0.028	-0.003	-0.050**	-0.036***	-0.019	-0.055***	-0.031*	-0.007	-0.050**
	(0.012)	(0.015)	(0.016)	(0.017)	(0.024)	(0.020)	(0.013)	(0.016)	(0.017)	(0.018)	(0.025)	(0.020)
Parental schooling years	0.077***	0.102***	0.049**	0.050***	0.066***	0.032	0.067***	0.090***	0.043**	0.045**	0.062**	0.027
	(0.016)	(0.023)	(0.022)	(0.017)	(0.025)	(0.024)	(0.016)	(0.024)	(0.022)	(0.017)	(0.025)	(0.024)
Father's occupation rank at child's age 14	0.002***	0.002***	0.002**	0.002***	0.002**	0.002**	0.002***	0.002**	0.002**	0.002***	0.002*	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father's party membership at child's age 14	0.360***	0.299	0.382**	0.304**	0.280	0.255	0.400***	0.357*	0.401**	0.324**	0.308	0.266
	(0.129)	(0.195)	(0.190)	(0.143)	(0.221)	(0.213)	(0.131)	(0.197)	(0.190)	(0.143)	(0.221)	(0.213)
Male	10.022***			10.118***			10.006***			10.101***		
	(0.096)			(0.111)			(0.095)			(0.110)		
Han Chinese	0.108	0.259	-0.005	0.230	0.208	0.365	0.097	0.244	-0.012	0.224	0.199	0.370
	(0.250)	(0.365)	(0.329)	(0.294)	(0.399)	(0.338)	(0.255)	(0.363)	(0.335)	(0.295)	(0.394)	(0.344)
Number of siblings	-0.054*	-0.054	-0.028	-0.108***	-0.108**	-0.070	0.002	0.012	0.002	-0.051	-0.053	-0.034
	(0.028)	(0.036)	(0.039)	(0.033)	(0.047)	(0.045)	(0.029)	(0.037)	(0.041)	(0.034)	(0.048)	(0.047)
Birth order index	0.061	0.034	0.053	0.046	0.051	-0.006	0.072*	0.054	0.058	0.043	0.050	-0.006
	(0.037)	(0.050)	(0.050)	(0.044)	(0.062)	(0.059)	(0.037)	(0.049)	(0.050)	(0.044)	(0.061)	(0.059)
Exposure age 0 to 3							1.455***	1.812***	0.814**	1.302***	1.207***	1.094***
							(0.218)	(0.322)	(0.319)	(0.249)	(0.408)	(0.366)
Number of siblings × Exposure age 0 to 3							-0.393***	-0.440***	-0.210**	-0.329***	-0.335***	-0.188*
							(0.068)	(0.108)	(0.095)	(0.069)	(0.113)	(0.096)
Observations	16051	8098	7953	12334	6165	6169	16051	8098	7953	12334	6165	6169
Adjusted R^2	0.560	0.225	0.164	0.569	0.235	0.159	0.561	0.228	0.164	0.570	0.236	0.160

Note:

1) Rural sample only

2) Robust Standard Errors (SEs) are presented in parentheses;

3) SEs are clustered at the county levels;

4) Unit: log of 10000 yuan; measured in 1990 price level

5) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;

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

Table A9: Main estimation including the famine and CR exposure variables

	Rural Total				Urban Total			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Agricultural output p.c.- age 0 to 3	0.194	0.176	0.193	0.175	-0.130*	-0.166***	-0.130*	-0.165***
	(0.172)	(0.180)	(0.170)	(0.178)	(0.073)	(0.061)	(0.072)	(0.061)
Log Industrial output p.c.- age 0 to 3	0.180**	0.193**	0.182**	0.194**	0.029	0.071	0.031	0.071
	(0.079)	(0.077)	(0.078)	(0.077)	(0.099)	(0.082)	(0.096)	(0.080)
Log Gov. Social expenditure p.c.- age 0 to 3	0.358***	0.353***	0.349***	0.345***	0.463***	0.473***	0.455***	0.466***
	(0.099)	(0.102)	(0.099)	(0.102)	(0.121)	(0.125)	(0.122)	(0.125)
Death rate (‰) - age 0 to 3	-0.032**	-0.027*	-0.036***	-0.032**	0.021	0.038	0.014	0.031
	(0.013)	(0.014)	(0.013)	(0.015)	(0.033)	(0.028)	(0.033)	(0.027)
Parental schooling years	0.071***	0.070***	0.067***	0.066***	0.114***	0.103***	0.112***	0.101***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.031)	(0.033)	(0.031)	(0.033)
Father's occupation rank at child's age 14	0.002***	0.002***	0.002***	0.002***	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father's party membership at child's age 14	0.386***	0.388***	0.400***	0.401***	0.002	0.023	0.003	0.023
	(0.130)	(0.130)	(0.131)	(0.131)	(0.307)	(0.300)	(0.307)	(0.300)
Male	10.024***	10.025***	10.006***	10.008***	11.324***	11.320***	11.322***	11.318***
	(0.096)	(0.097)	(0.095)	(0.095)	(0.238)	(0.234)	(0.238)	(0.234)
Han Chinese	0.110	0.110	0.097	0.096	-0.384	-0.403	-0.405	-0.421
	(0.252)	(0.251)	(0.255)	(0.254)	(0.827)	(0.816)	(0.822)	(0.812)
Number of siblings	-0.040	-0.041	0.002	0.001	-0.167**	-0.169**	-0.143**	-0.148**
	(0.028)	(0.028)	(0.029)	(0.029)	(0.067)	(0.067)	(0.068)	(0.068)
Birth order index	0.072*	0.069*	0.072*	0.070*	-0.030	-0.038	-0.031	-0.039
	(0.037)	(0.037)	(0.037)	(0.037)	(0.118)	(0.120)	(0.117)	(0.120)
Exposure age 0 to 3	0.738***	0.796***	1.455***	1.502***	1.731***	1.763***	1.934***	1.946***
	(0.164)	(0.171)	(0.218)	(0.222)	(0.308)	(0.327)	(0.329)	(0.345)
Exposure age 0 to 3 × Number of siblings			-0.393***	-0.390***			-0.391*	-0.357
			(0.068)	(0.068)			(0.230)	(0.230)
Exposure to Famine age 0 to 3		-0.076		-0.074		-0.680		-0.661
		(0.255)		(0.256)		(0.486)		(0.487)
Exposure to CR age 0 to 3		0.182		0.166		0.485		0.465
		(0.121)		(0.121)		(0.323)		(0.323)
Observations	16051	16051	16051	16051	2457	2457	2457	2457
Adjusted R^2	0.560	0.560	0.561	0.561	0.613	0.613	0.613	0.614

Note:

1) SEs are clustered at the county and prefecture levels for the rural and urban samples;

2) Unit: log of 10000 yuan; measured in 1990 price level

3) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;

4) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$;

Table A10: Main estimation considering different height growth period

	Rural sample			Urban sample		
	Local socioeconomic condition variables are the weighted average of					
	Age 0-3	Age 0-5	Age 0-7	Age 0-3	Age 0-5	Age 0-7
Log Agricultural output p.c.	0.427**	0.437**	0.464**	0.039	0.089	0.101
	(0.168)	(0.186)	(0.197)	(0.106)	(0.096)	(0.088)
Log Industrial output p.c.	0.206***	0.187**	0.172*	0.066	0.061	0.056
	(0.077)	(0.083)	(0.088)	(0.089)	(0.089)	(0.091)
Log Gov. Social expenditure p.c.	0.477***	0.514***	0.515***	0.893***	0.956***	1.011***
	(0.099)	(0.109)	(0.116)	(0.092)	(0.093)	(0.097)
Death rate (‰)	-0.032**	-0.039***	-0.047***	0.019	0.016	0.022
	(0.012)	(0.014)	(0.017)	(0.037)	(0.043)	(0.047)
Parental schooling years	0.077***	0.071***	0.068***	0.127***	0.120***	0.115***
	(0.016)	(0.016)	(0.016)	(0.033)	(0.033)	(0.033)
Father's occupation rank at child's age 14	0.002***	0.002***	0.002***	0.000	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father's party membership at child's age 14	0.360***	0.375***	0.380***	-0.093	-0.071	-0.052
	(0.129)	(0.130)	(0.130)	(0.294)	(0.293)	(0.295)
Male	10.022***	10.028***	10.028***	11.340***	11.344***	11.340***
	(0.096)	(0.096)	(0.096)	(0.245)	(0.243)	(0.240)
Han Chinese	0.108	0.118	0.119	-0.459	-0.429	-0.427
	(0.250)	(0.250)	(0.250)	(0.848)	(0.844)	(0.842)
Number of siblings	-0.054*	-0.045	-0.039	-0.226***	-0.205***	-0.194***
	(0.028)	(0.029)	(0.029)	(0.069)	(0.068)	(0.068)
Birth order index	0.061	0.057	0.054	-0.058	-0.070	-0.072
	(0.037)	(0.037)	(0.037)	(0.121)	(0.122)	(0.122)
Observations	16051	16051	16051	2457	2457	2457
Adjusted R^2	0.560	0.560	0.561	0.610	0.610	0.611

Note:
1) Robust Standard Errors (SEs) are presented in parentheses;
2) SEs are clustered at the county and prefecture levels for the rural and urban samples;
3) Unit: log of 10000 yuan; measured in 1990 price level
4) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;
5) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A11: Main estimation considering different height growth period

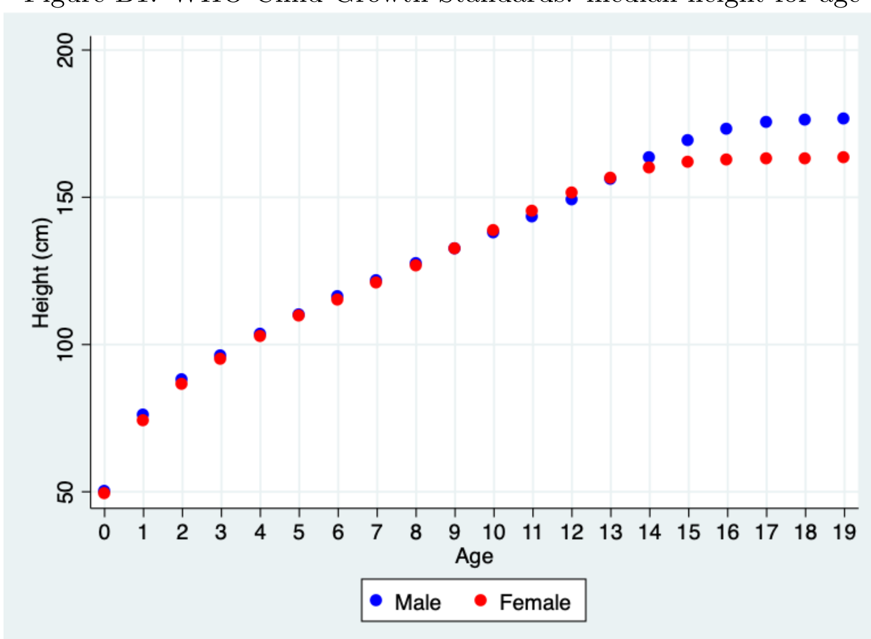
	Rural sample			Urban sample		
	Local socioeconomic condition variables are the weighted average of					
	Age 0-3	Age 0-5	Age 0-7	Age 0-3	Age 0-5	Age 0-7
Log Agricultural output p.c.	0.193	0.238	0.290	-0.130*	-0.080	-0.061
	(0.170)	(0.191)	(0.203)	(0.072)	(0.069)	(0.069)
Log Industrial output p.c.	0.182**	0.168**	0.156*	0.031	0.030	0.030
	(0.078)	(0.083)	(0.087)	(0.096)	(0.097)	(0.100)
Log Gov. Social expenditure p.c.	0.349***	0.392***	0.392***	0.455***	0.497***	0.537***
	(0.099)	(0.109)	(0.116)	(0.122)	(0.130)	(0.150)
Death rate (‰)	-0.036***	-0.047***	-0.058***	0.014	0.007	0.008
	(0.013)	(0.015)	(0.018)	(0.033)	(0.039)	(0.042)
Parental schooling years	0.067***	0.063***	0.061***	0.112***	0.109***	0.107***
	(0.016)	(0.016)	(0.016)	(0.031)	(0.031)	(0.031)
Father's occupation rank at child's age 14	0.002***	0.002***	0.002***	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father's party membership at child's age 14	0.400***	0.405***	0.405***	0.003	0.006	0.010
	(0.131)	(0.131)	(0.131)	(0.307)	(0.306)	(0.307)
Male	10.006***	10.010***	10.010***	11.322***	11.323***	11.322***
	(0.095)	(0.095)	(0.095)	(0.238)	(0.238)	(0.236)
Han Chinese	0.097	0.105	0.104	-0.405	-0.384	-0.383
	(0.255)	(0.254)	(0.255)	(0.822)	(0.823)	(0.824)
Number of siblings	0.002	0.007	0.011	-0.143**	-0.134*	-0.132*
	(0.029)	(0.029)	(0.029)	(0.068)	(0.068)	(0.068)
Exposure age 0 to 3	1.455***	1.333***	1.279***	1.934***	1.846***	1.774***
	(0.218)	(0.226)	(0.231)	(0.329)	(0.345)	(0.378)
Exposure age 0 to 3 × Number of siblings	-0.393***	-0.399***	-0.409***	-0.391*	-0.401*	-0.401*
	(0.068)	(0.068)	(0.069)	(0.230)	(0.232)	(0.232)
Birth order index	0.072*	0.066*	0.061*	-0.031	-0.043	-0.045
	(0.037)	(0.037)	(0.037)	(0.117)	(0.118)	(0.119)
Observations	16051	16051	16051	2457	2457	2457
Adjusted R^2	0.561	0.561	0.562	0.613	0.613	0.613

Note:
1) Robust Standard Errors (SEs) are presented in parentheses;
2) SEs are clustered at the county and prefecture levels for the rural and urban samples;
3) Unit: log of 10000 yuan; measured in 1990 price level
4) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;
5) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix B: How the damping coefficients and the exposure variable are generated

The WHO Multicentre Growth Reference Study (MGRS) provides reference growth standards for infants and young children using longitudinal surveys. The sample used consists of about 8,500 children from six countries - Brazil, Ghana, India, Norway, Oman, and the United States (Onis et al., 2004).[40] Figure B1 plots median height-for-age from the WHO Child Growth Standards. At age 0, the median height-for-age is about 50 cm for both females and males, and it jumps to 75 cm (female 74.005 cm; male 75.739 cm) at age 1, then further to about 87 cm (female 86.401 cm; male 87.802 cm) at age 2. As childhood evolves, the increment between two consecutive ages becomes smaller. At age 15, the growth for females stops, whereas for males it stops at around age 16. The height grows the most during the first two years of children's life.

Figure B1: WHO Child Growth Standards: median height-for-age



Using this growth reference data, we compute two variables: 1). the damping coefficient ω_n in equation (2) and 2). the exposure variable, which captures the share of an individual's first three years of life (age 0-3) that was at and after 1979, the year of the start of the economic reform and introduction of the OCP.

The basic idea is to calculate the relative increase in height between age 0 and 3. We calculate

[40] Data source: Boys 0-5 years old - https://www.who.int/childgrowth/standards/h_f_a_tables_z_boys/en/; Girls 0-5 years old - https://www.who.int/childgrowth/standards/h_f_a_tables_z_girls/en/; 5-19 years old - https://www.who.int/growthref/who2007_height_for_age/en/

an increase in height between two consecutive ages then divide it by the total increment between age 0 and 3. For example, female median height at age 0 is 49.148 cm, 74.005 cm at age 1, 86.401 cm at age 2, and 95.034 cm at age 3. The increments are 24.857 cm between age 0 and 1, 12.396 cm between age 1 and 2, and 8.633 cm between age 2 and 3. Since the total increment between age 0 and 3 is 45.886 cm (=95.034-49.148), each increment accounts for 54.17%, 27.01%, and 18.81% of the total increment between age 0 and 1, 1 and 2, and 2 and 3, respectively. The increment for each age is plotted in Figure B2.

This relative increase is the damping coefficient ω_n. Suppose one was born in 1970, and the log of agricultural output per capita is 0.1 in 1970, 0.2 in 1971, and 0.3 in 1972. Then $\omega_n E_{cn} = 0.1 * 0.5417 + 0.2 * 0.2701 + 0.3 * 0.1881$ in her case.

We apply this method to compute the exposure variable, which captures the degree of which one was exposed to the economic reform and introduction of the OCP between age 0 and 3. For example, if one was born in 1978, her exposure value is equal to 27.01%+18.81% as she was exposed to the economic reform and introduction of the OCP at age 1 and 2. If one was born in 1975, her exposure value is equal to zero as she did not experience those events before age 3. For those born in 1979 onwards, the exposure value is 100%. Figure B3 shows the distribution of the exposure variable for males (blue dots) and females (red dots) separately.

Figure B2: Increments between consecutive ages

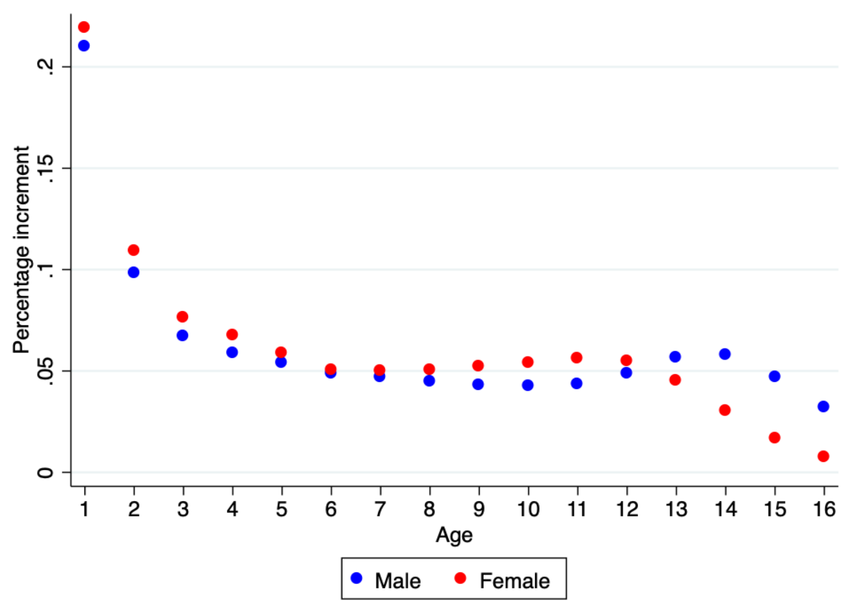
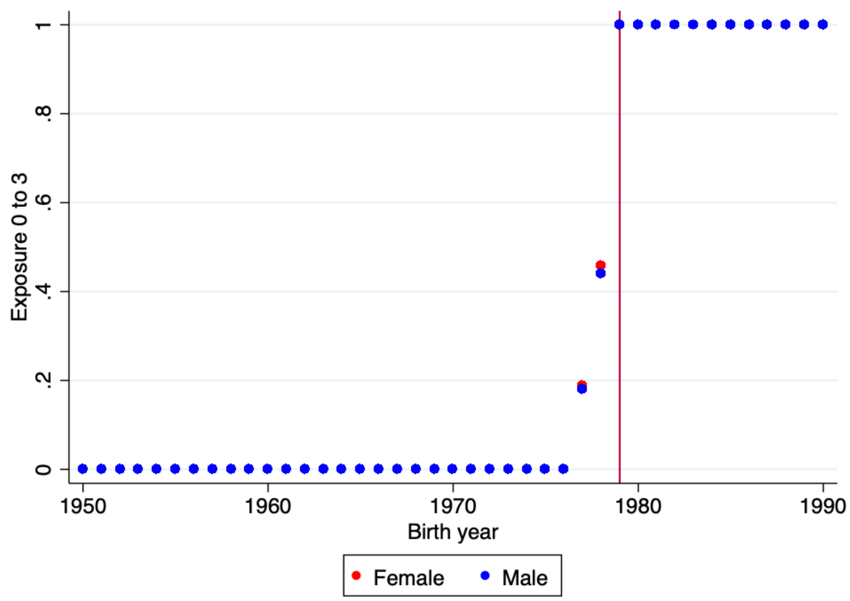


Figure B3: Cohort distribution of the 'exposure' variable



Appendix C: Local gazetteer data collection

C-1. Variable description

We collected annual data on the following variables from county gazetteers for the period between 1950 and 1990.

1) Industrial output value: This variable is the value of annual total industrial output produced in a county. Some counties use the corresponding year price, but others use a constant price across years. We convert all the values into the 1990 prices using a provincial-level retail price index. The unit is ten thousand yuan.

2) Agricultural output value: This variable is the value of total agricultural output produced in a county. It includes farming, forestry, animal husbandry, and fishery. Some counties use the corresponding year price, but others use a constant price across years. We convert all the values into the 1990 prices using a provincial-level retail price index. The unit is ten thousand yuan.

3) Local government expenditure on social services: This variable is the amount of annual local government expenditures on social services, including health care, education, social security benefit, culture, and science. We convert all the values into the 1990 prices using a provincial-level retail price index. The unit is ten thousand yuan.

4) Death rate: The definition of death rate is the number of deaths divided by the total population for each year. The unit is ‰.

5) Rural *hukou* population: It is the number of people whose household registration (*hukou*) is agricultural in a county. We generate the rural and urban *hukou* share variables using this variable and county total population information.

6) Population: We generate per capita output and expenditure variables using total population data. The unit is persons.

C-2. Data source

Table C1 provides data sources. The main sources are local gazetteers and statistical yearbooks. China has changed its administrative divisions significantly over the last decades, so some regions have multiple gazetteers. We do not merge data from different sources unless available observations for the overlapping periods are the same across the books.

C-3. Missing information

Table C2 shows data source and the proportion of missing values for each variable. We collected 6492 agricultural output observations (by county and year) from the local gazetteers and 1988 observations from the statistical yearbooks. 47.2% of the total observations of agricultural output and 42.6% of industrial output observations are predicted owing to missing information. Local government social expenditure information is mostly from the local gazetteers and 25.0% of the total observations in our sample are predicted. About half of the collected mortality rate observations are from the gazetteers and the other half is mainly collected from provincial-level population statistical yearbooks. Its proportion of missing observations is the lowest among the four county-level variables.

C-4. Data prediction

To fill missing values, we impute them using available county-level data and provincial-level data. Provincial-level data are from (NBS, 1999, 2010). The following equations (5) and (6) describe the first imputation methodologies.

$$\hat{x}_{c,p,t} = x_{c,t-1} \frac{x_{p,t}}{x_{p,t-1}} \tag{5}$$

where $\hat{x}_{c,p,t}$ is a predicted value for a missing observation of county c in province p in year t. $x_{c,t-1}$ is a non-missing observation of the same county c in year $t-1$. The observations for province p are available for both years t and $t-1$. First, we compute the ratio between $x_{p,t}$ and $x_{p,t-1}$ at provincial level. Second, we multiply $x_{c,t-1}$ by this ratio and compute $\hat{x}_{c,p,t}$. This methodology requires an assumption that the growth rate of x between years t and $t-1$ at provincial level is the same as that of the county c. We use this method for all county-level variables included in equation (2) (agricultural and industrial outputs, local government expenditure on social services, and death rate).

There are cases where the gazetteers or yearbooks provide county-level gross domestic product

Table C1: Data source

Province	Source
Anhui	1) Local gazetteers
Beijing	1) Local gazetteers 2) Comprehensive Statistical Data and Materials on 50 Years of New China
Chongqing	1) Local gazetteers
Fujian	1) Local gazetteers 2) Fujian Statistical Yearbook "New China 60 Years-Fujian"
Gansu	1) Local gazetteers 2) Gansu Population Statistical Yearbook 1949-1987;
	3) Gansu Statistical Yearbook "New China 60 Years - Gansu"
Guangdong	1) Local gazetteers 2) 2 City Statistical Yearbooks
Guangxi	1) Local gazetteers 2) Guangxi Population Statistical Yearbook 1949-1985
Guizhou	1) Local gazetteers 2) Guizhou Population Statistical Yearbook 1949-1984
	3) Guizhou Statistical Yearbook "Guizhou 60 Years 1949-2009"
	4) Guizhou Statistical Yearbook "Guizhou 30 Year" 5) City Statistical Yearbook
Hebei	1) Local gazetteers 2) Hebei Statistical Yearbook "New Hebei 60 Years"
	3) Hebei Statistical Yearbook "New Hebei 50 Years"
	4) Hebei Province Population Statistical Yearbook 1949-1984
Heilongjiang	1) Local gazetteers 2) Heilongjiang Statistical Yearbook "Longjiang 60 Years"
	3) Heilongjiang Population Statistical Yearbook
Henan	1) Local gazetteers 2) Henan Statistical Yearbook "Henan 30 Years"
Hubei	1) Local gazetteers 2) Hubei Statistical Yearbook 1985 and 1987
	3) Hubei Statistical Yearbook "Hubei 30 Years"
Hunan	1) Local gazetteers 2) Hunan Population Statistical Yearbook 1949-1991
	3) City statistical yearbook
Jiangsu	1) Local gazetteers 2) Jiangsu Population Statistical Yearbook 1949-1985;
	3) Jiangsu Statistical Yearbook - "Jiangsu 60 Years"
Jiangxi	1) Local gazetteers 2) Jiangxi Statistical Yearbook "New China 50 Years - Jiangxi";
	3) Jiangxi Population Statistical Yearbook 1949-1985
Jilin	1) Local gazetteers 2) Jilin Population Statistical Yearbook 1949-1984
Liaoning	1) Local gazetteers 2) Liaoning Statistical Yearbook "Liaoning 60 Years"
	3) Liaoning Statistical Yearbook "Liaoning 40 Years" 4) Liaoning Population yearbook 1949-1984
	5) city statistical yearbook
Shaanxi	1) Local gazetteers
Shandong	1) Local gazetteers 2) Shandong Statistical Yearbook "New China 50 years - Shandong province"
	3) Shandong Population Statistical Yearbook 1949-1984 4) City economic statistical yearbook
Shanghai	1) Local gazetteers 2) Shanghai Population Statistical Yearbook 1949-2000
	3) Shanghai Statistical Yearbook - "Shanghai 60 Years of Statistics compilation"
Shanxi	1) Local gazetteers 2) Shanxi Population Statistical Yearbook
	3) Shanxi Statistical Yearbook "Shanxi 60 Years"
Sichuan	1) Local gazetteers 2) Sichuan Statistical Yearbook "Sichuan Statistics Compilation (1979-1990)"
Tianjin	1) Local gazetteers 2) Tianjin Population Statistical Yearbook 1979-1988
Yunnan	1) Local gazetteers 2) Yunnan Population Statistical Yearbook 1949-1988
Zhejiang	1) Local gazetteers 2) Zhejiang Statistical Yearbook "Zhejiang 60 Years"
	3) Zhejiang Population Statistical Yearbook 1949-1985
All provinces	China City Statistical Yearbook 1985, 1986, 1987, 1988, 1989, and 1990
	Retail price index: Comprehensive Statistical Data and Materials on 50 Years of New China

Table C2: Data source and the proportion of missing values

	Data source		**Missing Obs.**	**Total Obs.**	
	Gazetteer	Yearbook	(a)	(b)	(a/b)
Rural sample					
Agricultural output	6492	1988	7571	16051	0.472
Industrial output	7219	1990	6842	16051	0.426
Government expenditure	11758	258	4008	16051	0.250
Death rate	6889	6573	2589	16051	0.161
Urban sample					
Agricultural output	293	1359	805	2457	0.328
Industrial output	364	1214	879	2457	0.358
Government expenditure	248	1959	250	2457	0.102
Death rate	226	1731	500	2457	0.204

Column 1 shows the number of observations collected from local gazetteers, column 2 shows the number of observations collected from yearbooks, column 3 indicates the number of missing values.

(GDP) data, which covers the whole 40-year period, while economic output information is not fully available. In this case, we impute missing values of the output variables using county-level GDP data as shown in equation (6):

$$\hat{x}_{c,p,t} = gdp_{c,t} \frac{x_{p,t}}{gdp_{p,t}} \tag{6}$$

where $\hat{x}_{c,p,t}$ is a predicted value of a missing observation of county c in provincial p at year t. Available observations are, 1) $gdp_{c,t}$: GDP for county c in year t, 2) $x_{p,t}$: economic output for province p in year t, and 3) $gdp_{p,t}$: GDP for province p in year t. First, we compute the ratio between GDP and economic output levels at the provincial level. Then we multiply this ratio by the county-level GDP $gdp_{c,t}$ and calculate $\hat{x}_{c,p,t}$. This methodology requires an assumption that the ratio between GDP and economic output values is the same between provincial-level and county-level data. More details on imputation methodologies are available upon request from the authors.

To test the sensitivity of this imputation method, we also impute missing values using the following prediction models:

$$e_{cpt} = \alpha + \beta e_{pt} + u_{cpt} \tag{7}$$

$$e_{cpt} = \alpha + \beta e_{pt} + \delta_t + u_{cpt} \tag{8}$$

$$e_{cpt} = \alpha + \beta e_{pt} + \sum_{p=1}^{24} \gamma_p e_{pt} + u_{cpt} \tag{9}$$

where e_{cpt} is a county-level variable of county c in province p at year t and e_{pt} is a provincial-level variable. β captures the association between e_{cpt} and e_{pt}. Equation (7) includes e_{pt} and equation (8) adds year fixed effects δ_t. To consider the potential heterogenous associations between e_{cpt} and e_{pt} across provinces, which is captured by γ_p, equation (9) includes interactions between provincial dummies and e_{pt}. We estimate α, β, δ, γ_p then use them to predict missing values of e_{cpt}. Table C3 presents the estimation results for equation (2) using predicted missing values based on these three prediction models. As can be seen, changing imputation methods do not alter our results qualitatively.

Table C3: Estimated results for equation (2) using different imputation methods for missing values

	Rural sample				Urban sample			
		Prediction methods				Prediction methods		
	Main result	Eq.(7)	Eq. (8)	Eq. (9)	Main result	Eq. (7)	Eq. (8)	Eq. (9)
Log Agricultural output p.c.- age 0 to 3	0.427**	0.632***	0.332**	0.833***	0.039	-0.036	0.031	0.069
	(0.168)	(0.155)	(0.139)	(0.173)	(0.106)	(0.134)	(0.112)	(0.157)
Log Industrial output p.c.- age 0 to 3	0.206***	0.141**	0.057	0.069**	0.066	0.184**	-0.031	-0.097
	(0.077)	(0.059)	(0.055)	(0.030)	(0.089)	(0.090)	(0.124)	(0.077)
Log Gov. Social expenditure p.c.- age 0 to 3	0.477***	0.311***	0.443***	0.267***	0.893***	0.820***	1.014***	0.881***
	(0.099)	(0.056)	(0.069)	(0.053)	(0.092)	(0.144)	(0.112)	(0.139)
Death rate (‰) - age 0 to 3	-0.032**	-0.035***	-0.043***	-0.038***	0.019	0.010	0.005	-0.005
	(0.012)	(0.013)	(0.013)	(0.012)	(0.037)	(0.037)	(0.036)	(0.039)
Parental schooling years	0.077***	0.087***	0.094***	0.087***	0.127***	0.133***	0.125***	0.133***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.033)	(0.031)	(0.031)	(0.031)
Father's occupation rank at child's age 14	0.002***	0.002***	0.002***	0.002***	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father's party membership at child's age 14	0.360***	0.345***	0.322**	0.333**	-0.093	-0.107	-0.054	-0.107
	(0.129)	(0.129)	(0.128)	(0.129)	(0.294)	(0.296)	(0.290)	(0.299)
Male	10.022***	10.020***	10.016***	10.020***	11.340***	11.315***	11.314***	11.305***
	(0.096)	(0.096)	(0.097)	(0.096)	(0.245)	(0.241)	(0.238)	(0.245)
Han Chinese	0.108	0.089	0.087	0.096	-0.459	-0.445	-0.407	-0.481
	(0.250)	(0.251)	(0.250)	(0.253)	(0.848)	(0.854)	(0.864)	(0.843)
Number of siblings	-0.054*	-0.067**	-0.083***	-0.069**	-0.226***	-0.245***	-0.233***	-0.256***
	(0.028)	(0.028)	(0.028)	(0.029)	(0.069)	(0.067)	(0.065)	(0.067)
Birth order index	0.061	0.064*	0.070*	0.065*	-0.058	-0.051	-0.043	-0.047
	(0.037)	(0.037)	(0.037)	(0.037)	(0.121)	(0.122)	(0.121)	(0.121)
Observations	16051	16051	16051	16051	2457	2457	2457	2457
Adjusted R^2	0.560	0.559	0.559	0.559	0.610	0.609	0.611	0.609

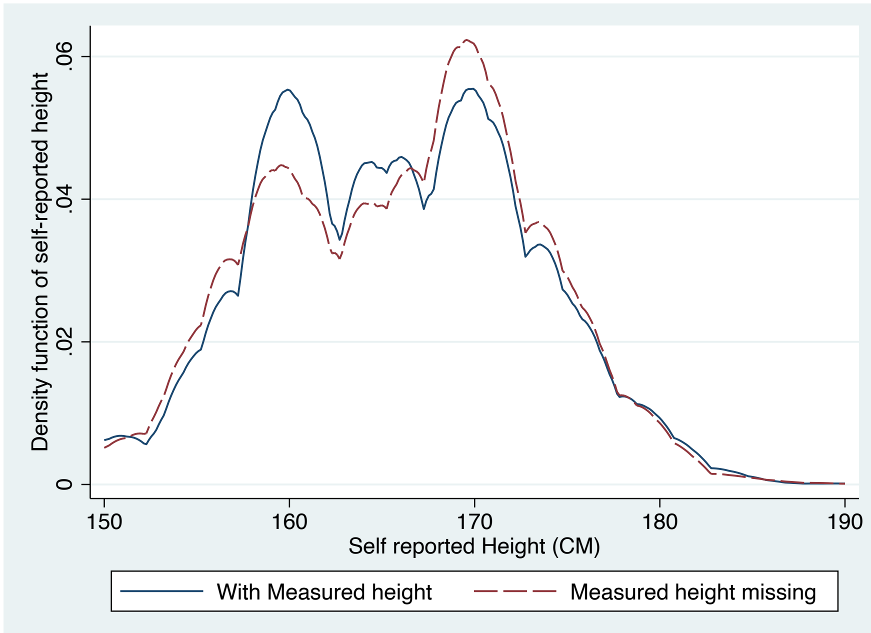
Note:
1) Robust Standard Errors (SEs) are presented in parentheses;
2) SEs are clustered at the county and prefecture levels for the rural and urban samples;
3) Unit: log of 10000 yuan; measured in 1990 price level
4) County fixed effects, a share of urban population in each county, and a group of dummy variables indicating the missing values for father's party membership and occupation as well as birth order variables are included;
5) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix D: CHNS/RUMiC sample construction

The China Health and Nutrition Survey (CHNS) is a longitudinal survey of households and individuals conducted by the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health at the Chinese Center for Disease Control and Prevention since 1989. We use all rounds between 1989 and 2011 (1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009, and 2011), and the survey covers 9 provinces in China - Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi, and Guizhou. There are 157,286 observations (including repeated individuals) between the 1989 and 2011 waves, and 82,122 of the individuals of them were born between 1950 and 1990 and aged between 20 and 60 at the survey. Birth region information is a crucial part of our analysis, but the CHNS did not collect individual birth county information. Instead, it provides birth province of each household's head and her/his spouse but only 7,827 out of 82,122 observations have that information. We assume that all household members were born in the same province as the household head, and we use the spouse birth province instead when the household head's birth province information is missing. As a result, 73,789 (18,992 individuals) out of 82,122 observations have birth province information. Among them height observations are available for the 14,993 individuals. To minimise the measurement error due to either shrinkage or premature height, we use the maximum height of available observations for each individual. We include 12,348 (82.3%) of 14,993 individuals who were born in the 9 CHNS survey provinces. The CHNS does not collect information on the *hukou* status at birth so we instead use the current *hukou* status to define the rural sample. This assumption is reasonable considering that 85% of individuals in the CFPS rural sample currently hold the rural *hukou* in 2010. The final CHNS rural sample consists of 7,044 individuals.

The Rural-Urban Migration in China (RUMiC) project was established to investigate the impacts of internal migration in China by a group of researchers at the Australian National University. The survey began in 2008 and ended in 2016. The measured height was available between 2009 and 2011. Due to the high attrition rate, we only used the 2009 data. The RUMiC survey sampled rural-to-urban migrants in 15 destination cities. However, as they are migrants with rural *hukou*, their household registration locations are the sending province, which indicates that RUMiC 2009 sample came from 28 out of 30 provinces. In the 2009 survey there are 7418 individuals who were born between 1950 and 1990. Of those, 3738 observations have non-missing values for measured the height variable, whereas all 7418 observations self-reported the height variable. Figure D1 presents the distribution of self-reported height for the samples with and without missing values on measured height. As can be seen from the figure that the two distributions are quite similar. The mean difference between the two sample on self-reported height is 0.24cm with the t-statistic being 1.44.

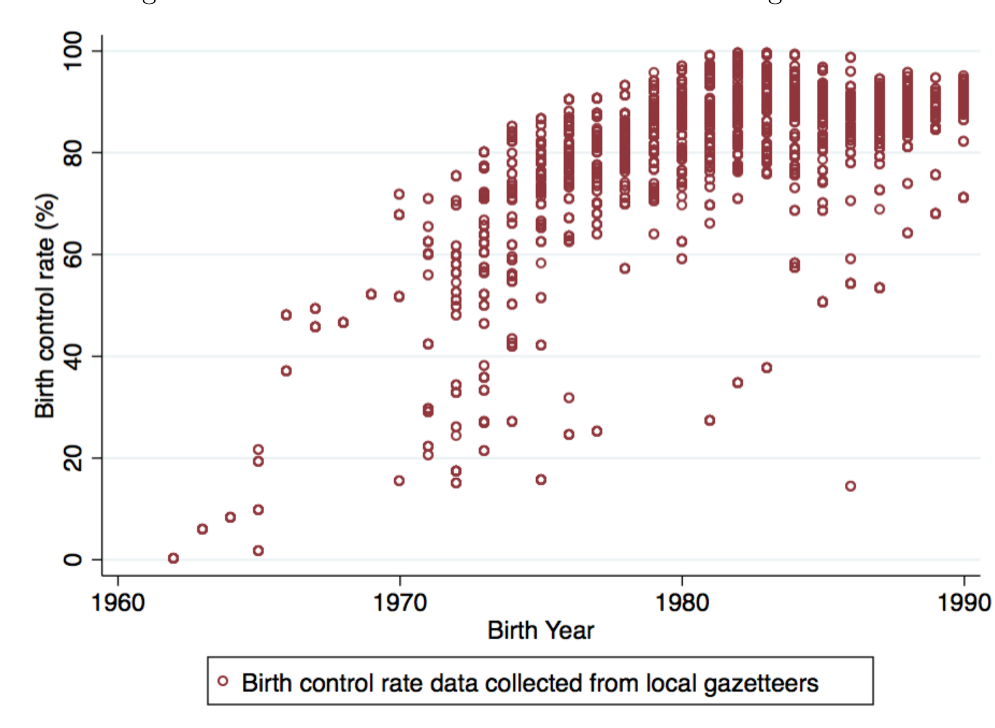
Figure D1: Comparison of reported height between missing and non-missing measured height samples: RUMiC 2009



Appendix E: Instrumental variable construction

We use an instrument for the endogenous variable 'number of siblings' constructed by Chae (2020), which is the proportion of fertile women who use any contraceptive method in their birth county for the period around the individual's birth. This variable is called '*jieyu*' in Chinese, and we translate this as 'Birth Control Rate (BCR)'. It is reported in most regional gazetteers and used to reflect each local government's fertility control intensity.

Figure E1: Birth Control Rate collected from local gazetteers



There is a total of 159 counties in our main sample, and we have the BCR information for 103 counties (64.8%) as plotted in Figure E1, of which 59.22% are from county-level gazetteers and 29.21% from prefecture-level gazetteers. In the remaining counties, we are unable to find gazetteer information. Further details of how to impute missing values are available in Chae (2020).

Consistent with Chae (2020), we use the average rate over the period of three years before and three years after the individual's birth which is $\bar{B}_{ct}$ in equation (10).[41] $\bar{B}_{ct}$ captures the family planning policy intensity a few years around the timing of one's birth, which should highly correlated with the number of siblings one has. More specifically:

[41] Chae (2020) conducts a sensitivity test of varying years of the coverage including 2 years before and after (=within 5 years of one's birth) and 1 years before and after (=within, and 2 year before and after).

$$\bar{B}_{ct} = \frac{1}{7}\sum_{t=-3}^{+3}\frac{CW_{ct}}{W_{ct}} \tag{10}$$

where $\bar{B}_{ct}$ is the average BCR within 7 years of one's birth year t in birth county c. CW_{ct} is the number of fertile women who either are sterile or use contraception and W_{ct} is the total number of fertile women in county c in year t. We average the proportion of CW_{ct} to W_{ct} between 3 years before and after one's birth including one's birth year. In other words, we use 7 year average of $\frac{CW_{ct}}{W_{ct}}$ for $t-3$, $t-2$, $t-1$, t, $t+1$, $t+2$, and $t+3$ as the instrument for the number of siblings of an individual born in t. The contraceptive method includes (1) female sterilisation, (2) male sterilisation, (3) intrauterine contraceptive devices, (4) oral contraceptive pills, (5) abortions, (6) pills in external use, (7) injections, (8) condoms, and (9) other medicine or medical devices. Intuitively, $\bar{B}_{ct}$ captures the likelihood of having siblings within 3 years before or after one was born. Figure E2 plots the distribution of the Birth Control Rate. Further details of how to construct the instrument are available in Chae (2020).

Figure E2: The instrument - Birth Control Rate (BCR) (%)