Long-term Effect of *In Utero* Exposure to Land Reform on Academic Performance in China

Douglas Almond, Columbia University & NBER
Hongbin Li, Tsinghua University
Binzhen Wu, Tsinghua University
Shuang Zhang, Cornell University

June 9, 2012

Preliminary; Please Do Not Circulate

Abstract

While a growing body of literature exploits extreme shocks *in utero* to test the fetal origins hypothesis, empirical evidence on the later-life impacts of income shock is relatively thin. This paper examines the effect of prenatal exposure to the 1978-84 land reform in China on academic performance, as captured by college entrance exam scores. By replacing collectivized farming with more autonomous household farming, the economic liberalization is widely documented to have increased rural household income and reduced poverty. Using each test-taker's year of birth and county in infancy matched to the year land reform started in his/her county in infancy for 1068 counties, we find that high school students born just after the first post-reform harvest perform better on college entrance exams, math especially, and are more likely to be admitted by first-tier and top-ranked colleges. These effects are substantially stronger for boys. Our findings suggest that income might be beneficially targeted to the prenatal period and that "pro-growth" policies can lay the foundation for human capital accumulation.

1 Introduction

The timing of investments during childhood can be crucial to later life outcomes (Heckman, 2007). While a growing body of the "fetal origins" literature explores the later-life impacts of extreme shocks, empirical evidence on early-life exposure to public policies affecting economic wellbeing is relatively thin (Almond and Currie, 2011b). In this paper, we examine the long-term effect of *in utero* exposure to "the world's largest antipoverty program" (McMillan, 2002): the pro-market land reform in China that occurred between 1978 and 1984 following the death of Mao.

We exploit the fact that the land reform decision was taken at a local level using new data on the year reform started by county. For 1,068 counties we link microdata on cohorts born around the time of reform in their college entrance exam scores between 1999 and 2001, nearly 2 million testing records. Ours is the first paper to consider long-term developmental effects of such a sweeping reform. Despite recognition that the 1978-84 land reform catalyzed China's transition to a market economy, previous studies have been confined to short-term productivity and income growth (McMillan et al., 1989; Lin, 1992). Little is known about broader impacts of the reform on other dimensions of human wellbeing, such as health and education, particularly at the individual level. Furthermore, our identification strategy departs from previous analyses that relied on the variation among 28 provinces in the reform diffusion (Lin, 1992).

First, we confirm the "first stage" effects of province-level studies and find prompt gains in agricultural output at the county level, first observed one year after reform (of around 5%, increasing to over 10% three or more years after reform). Turning to the 1999-2001 test records, we find that cohorts of high school students born two years after the reform perform better on college entrance exams, particularly on math. They are also more likely to be admitted by first-tier and top ranked colleges. The impact estimates are relatively modest in magnitude: math test scores increase by just under 1% and attending a selective college by 4-5%. Interestingly, these effects are substantially stronger for boys.

Our paper contributes to two literatures. First, it expands the fetal origins literature to examine the long-term effect of *in utero* income shock induced by public policy on human capital, and thus helps inform the generalizability and policy relevance of the fetal origins effects. Second, we provide evidence and an analytical framework for interpreting the gender difference in the fetal origins effects as observed in many studies (e.g., Van Den Berg et al., 2006; Field et al., 2009; Maccini and Yang 2009; Almond et al, 2010; Bhalotra and Venkataramani, 2011).[1] The epidemiological literature generally attributes empirical differences in fetal origins effects by gender to biological

[1] Van Den Berg et al., (2006) find that exposure to adverse economic condition at birth increase mortality rate for men but not for women. Bhalotra and Venkataramani (2011) find that men exposed to sulfa drugs in infancy benefit more in terms of schooling, employment and income in adulthood. Almond et al., (2010) find that men exposed to famine are more likely to be illiterate, less likely to work and get married. In contrast, Field et al., (2009) find that exposure to iodine supplementation *in utero* increases schooling for girls, but not for boys, in Tanzania. Maccini and Yang (2009) find larger influences of a rainfall shock in the birth year on women's health, education and economic outcomes later in life.

differences (e.g., Kraemer, 2000). The purely biological explanation, however, can not be easily reconciled with existing evidence, for example, the larger negative effects of rainfall shock at birth for women in Maccini and Yang (2009). On the other hand, recent work in economics on responsive investments shows that parents may reinforce early differences by investing more in better endowed children (Datar et al., 2010; Adhvaryu and Nyshadham, 2011).[2] In the context of a society with son preference, it is particularly important to consider how reinforcement of parents might generate the gender difference. Although we unfortunately do not have direct measures of parental investments, we argue that biology is unlikely to explain the gender difference in the effect of land reform. Instead, our evidence is consistent with differential responses to a given endowment shock depending on whether the child is a boy or girl.

Our difference-in-differences approach compares academic performance among individuals born before and after the first post-reform harvest between counties that reformed earlier and those that reformed later. A key identifying assumption is the absence of systematic differences in the pre-existing trends of academic performance across counties with different reform timing. We have included county-specific time trends that capture any differential patterns for test takers born prior to the first harvest across counties. Moreover, through comprehensive reading on the evolution of the reform policy and the previous literature at the province level, we also find out the primary determinants of the reform timing at the county level and account for their possibly time-varying effects.

Finally, we argue that our analysis is impervious to various plausible threats to identification and inference. There is limited scope for bias from endogenous migration: we observe and assign each test taker's county before age 1 in the microdata.[3,4] It is unlikely that other post-Mao reforms followed the timing of land reform closely by county and have had differential impacts for individuals born before and after the first harvest, except for the introduction of the One Child Policy (OCP). Our results are robust to inclusion of controls for exposure to the OCP. We test for, and rule out possible hypotheses about endogenous selection into the sample of test takers. For example, there is no evidence that land reform affects one's chances of going to high school or dropping out from high school, one's decisions to apply for college and take the college entrance exams, and one's age at the exams.

The remainder of the paper is organized as follows. Section 2 introduces background on the post-Mao land reform and describes the education system in China. Section 3 describes the

[2] Datar et al., (2010) show that low birth weight children are less likely to be breastfed, have fewer well-baby visits, are less likely to be immunized, and are less likely to attend preschool than normal birth weight siblings. Adhvaryu and Nyshadham (2011) find that parents reinforce the higher cognitive endowments of children who received in utero iodine supplementation, by investing more in vaccinations and early life nutrition.

[3] This stands in contrast to the United States, where large sample microdata generally do not report county of birth.

[4] Mobility of parents induced by demand for land between 1978 and 1984 was unlikely because land was allocated based on family size within one's village of birth, and internal migration had been under strict control under the *Hukou* system since the 1950s until its first relaxation in 1985 (Wang, 2005).

microdata for empirical analysis. Section 4 discusses our identification strategy. Main results are presented in Section 5, and results for robustness are in Section 6.

2 Background

2.1 The post-Mao land reform

China had pursued collective agriculture for more than two decades since the mid-1950s. Under this system, workers generally received daily fixed work points and got paid at the end of the year (Lin, 1988). The incentive to work was low and agricultural productivity was stagnant. From 1956 to 1977, there had been virtually no change in grain output per capita (Zweig, 1987). Following the death of Mao Zedong and the end of the Cultural Revolution, a decollectivization reform was initiated in 1978. Only except for land ownership, collective system was abandoned. Collectively owned land was contracted to individual households for up to 15 years.[5] Consequently, the basic decision-making unit was shifted from the collective farm to individual households, who could make their own input decisions and receive all the residual income from the land after meeting the tax and quota sales obligations to the state (Perkins, 1988; Sicular, 1991).

The HRS was initially worked out by farmers without approval from the central government (Lin, 1987). At the end of 1978, a small number of production teams in Anhui Province, which suffered from a severe drought in that year, start to experiment with contracting land and output quotas to individual households (Lin, 1987; Yang, 1996). The initial position of the Central Committee of the Chinese Communist Party (CCP) on household farming was in a strong opposition. The document "Regulations on the Management of Rural People's Commune" passed by the CCP in the November of 1978 clearly stipulated that contracting to individual households was not permitted.

The prohibition was relaxed in September 1979 by allowing a few exceptions to households living in areas that were peripheral, distant, mountainous, and isolated due to transportation difficulties.[6] In September 1980, the Central Document No.75 issued by the Central Committee further allowed poor and remote areas and production units heavily dependent on state subsidies to contract land and output quota to households. By August 1981, the Central Committee's position on household farming became more liberalized in a mission sent to fifteen provinces "contracting to households is not only a means of relieving poverty but also ways of enhancing productivity; and it hasn't changed the production relations of the collective economy".[7] In January 1982, the Central Document No.1 officially announced that "the HRS is the production responsibility system of the socialist economy", which first showed the CCP's commitment to popularize HRS.

[5] It was extended to 30 years in 1993.

[6] *Agriculture Yearbook of China 1980*, 1981, Beijing, Agricultural Press.

[7] People's Daily, August 4th, 1981

2.2 Variation in the reform timing

Data on the timing of land reform at the county level were previously undiscovered. We have collected new data on the year HRS was introduced by county from 1242 county gazetteers in which local history from 1949 to the 1980s was recorded. Specifically, it is the year when collectively owned land was first contracted to individual households in a few villages of each county; and it usually took 2-3 years to spread the HRS to the whole county. These counties represent two-thirds of all counties that have ever published gazetteers.[8] Because land reform occurred in rural areas, our sample includes locations that were rural counties at the time of the reform.[9] In Figure 1, we plot the fraction of counties that had introduced HRS between 1978 and 1984. Only 1.85 percent of counties were the reform pioneers in 1978. The vast majority reformed between 1979 and 1981, with the peak of 45 percent in 1980. By 1984, all counties had adopted HRS.

Unlike many top-down policy implementations, the 1978-84 land reform is a bottom-up process. The reform timing by county is therefore not randomly assigned. Fortunately, what drives the local timing has not been a black box. The evolution of the reform policies suggests two primary sources of the variation: drought and poverty prior to the reform. A severe drought led to large declines in agricultural production, which in turn provided the local government incentive to reform.[10] Particularly, the negative production shock spurred taking the political risk as reform starters, since contracting land to individual households were not permitted in earlier years. Poor and remote counties were among the first permitted to adopt HRS by the central government as a means to reduce national poverty rates.

Moreover, existing literature on HRS adoption at the province level provides three additional insights (Lin, 1987; Yang, 1996; Chung, 2000). First, the diffusion of HRS was faster where reduction in monitoring cost is higher and thus productivity gains are larger. Using size of production team to measure monitoring cost, previous studies show mixed findings.[11] The second hypothesis is that provinces that suffered more from the 1959-61 Famine reformed earlier, because the more painful the lesson from the Famine, the less likely the province to favor collective farming. Supportive evidence is provided by Yang (1996) and Bai and Kung (2011). Lastly, Yang (1996) argues that provinces further from Beijing had more freedom to initiate earlier reform.

We provide the first evidence on the correlation between land reform timing and potential determinants at the county level. By matching the county-level data on reform timing with county-by-year data on precipitation (see Data Appendix), we examine whether drought had led to land

[8] The other one-third of counties do not report the timing of HRS adoption, or report it as "the late 1970s" or "the early 1980s", which are not useful for our identification strategy.

[9] City districts are defined and excluded by using the county code in the 1982 Census and the official definition.

[10] Bai and Kung (2011) provide indirect evidence using province level data. They find that provinces that suffered more in the 1959-61 Famine started land reform earlier when struck by bad weather. The interpretation is that the Famine experience undermined the local belief that collective farming could effectively cope with negative weather shocks.

[11] Lin (1987) finds that provinces with larger production teams reformed earlier, while Chung (2000) has the opposite finding.

reform. Notably, land reform is an irreversible event, implying that drought prior to reform might affect the decision to reform, but drought after would not. In this test, we assign one to the first year of reform, zero before the first year and missing values after. As suggested by Chinese Academy of Agricultural Sciences (1984), the growth of rice, the top one grain output in China, largely depends on rainfall at the beginning of the growing season, which starts in March or April. If drought affected reform timing through its impact on crop yield, we would expect a stronger correlation regarding spring drought. In Table 1A, column 1 shows no correlation between the first year of reform and drought defined by average monthly precipitation in the whole growing season (March to September) in the reform year and the year before.[12] From column 2 to 5, we measure drought by monthly precipitation from March to June separately. Consistent with the expectations, droughts in March and April of the reform year and one year before increased the probability of earlier reform, which are all statistically significant at the one percent level. In contrast, droughts in May and June had no impact on the reform timing.

Next, to test the cross-sectional relationship between reform timing and other local characteristics as discussed above, we have collected and constructed county-level measures prior to the reform from various sources. Poverty is captured by grain output per capital in 1977 that are collected from county gazetteers, and remoteness is measured by distance to province capital using 1982 Census. Size of production team is proxied by the density of labor force (aged 16-60) in 1977.[13] Famine intensity is measured by the average birth cohort size in 1953-1957 divided by the average cohort size in 1959-1961 using the 1982 Census.[14] We also calculate the distance to Beijing to capture freedom in local policy-making. Table 1B reports both estimates from univariate regression for each determinant and those from multivariate regressions. Consistent with previous literature, counties that were initially poor, that had larger production teams in 1977 and higher famine intensity in 1959-1961, and further from the central government decided to reform earlier. Controlling for grain output per capita in 1977 forces us to drop two thirds of the sample due to lack of data. To get close to the full sample with data on reform timing and explore the robustness, we omit grain output in the last column and find robust results for labor force density and famine intensity, though it decreases the standard error of distance to Beijing that becomes statistically significant.[15]

Although we find that droughts in March and April played an important role in initiating the reform, droughts are not ideal instruments for reform timing because rainfall shocks have been found to affect educational outcomes directly (Maccini and Yang, 2009). To attribute changes in individual academic performance to prenatal exposure to land reform, it is crucial to account for

[12] Because the first month of reform are mostly not documented, a drought in the growing season is likely to affect reform at the second half of the current year or next year.

[13] Density is calculated by population size aged 16-60 in 1977 divided by area at the county level using 1982 Census.

[14] Meng et al. (2009) use a similar measure of famine intensity using the 1990 Census.

[15] Grain output per capita in 1977 is negatively correlated with distance to Beijing, and positively correlated with labor force density in 1977.

county features that contribute to the local decision to reform and that have time varying effects on test scores. In all regressions on the effect of land reform, we control for droughts in March and April in current year and the year before, as well as time-invariant determinants of reform timing by county interacted with time fixed effects.

2.3 Land reform and agricultural output

By linking rewards directly to effort, the HRS provided economic incentives to individual households (Sicular, 1991). The reform has been found to have greatly improved agricultural productivity and increased rural household income between 1978 and 1984. McMillan et al. (1989) suggest that over three-quarters of the productivity increase during this period could be attributed to the incentive effects of the HRS. Using provincial data on the diffusion of HRS, Lin (1992) also finds that the reform accounts for half of the output growth. The reform is also widely recognized for its achievements in lifting hundreds of millions of rural households out of poverty (World Bank, 2000).[16]

Grain production represents the vast majority of agricultural production from the 1970s to the mid 1980s in China. Using unique data on total population and total output of grain production by county collected from county gazetteers,[17] we provide the first quantitative evidence on the productivity gain of the 1978-84 land reform at the county level. There are 400 counties that report both the reform timing and the complete year-by-year total population and total output of grain production from 1974 to 1984, among which 382 counties are matched with data on county controls correlated with reform timing. Records on grain output in the 1970s are particularly valuable because county-level statistics have been released systematically only since the 1980s in China. They are also arguably reliable because these data were originally from local official archives (Xue, 2010).[18]

In Figure 2, we plot grain output per capita by the year relative to the first year of HRS adoption.[19] Time 0 indicates the first year of reform. The pattern prior to land reform is relatively flat, which is consistent with the literature in that agricultural productivity growth under the collectivized system was sluggish. There is an obvious jump of grain output one year after the first reform year, suggesting that the first post-reform harvest was in the second reform year. The rising pattern is persistent till four years after the reform. An important message from this "first-stage" figure is that the post-reform indicator should be assigned as 1 from the second year of HRS adoption.

[16]Official estimates indicate that rural poverty rate declined from 30 percent in 1978 to 5 percent in 1998 (World Bank, 2000).

[17]Grain crops generally include rice, wheat, corn and potato.

[18]Because the purpose of compiling county gazetteers is to accurately record local history, local historians working in the county gazetteer office have less incentive to manipulate the grain output data.

[19]Grain output per capita is measured by the total output of grain production divided by total population for each year.

We then estimate the effect of land reform on grain output per capita in a difference-in-difference framework and report results in Table 2. In column 1, county fixed effects, year fixed effects, and county specific time trends are controlled for. Panel A shows that HRS adoption increases grain output per capita by 3.2 percent, and Panel B presents monotonic increases in grain output from one year to four years after land reform. The largest output gain appears four years after the HRS adoption, which is a 16 percent increase. The increases in grain output are statistically insignificant five years after land reform.[20] In column 2, we further control for drought in March and April of the current year and one year before, and other determinants of reform timing interacted with time fixed effects. The point estimate in Panel A decreases to 2.6 percent and standard error remains the same, while estimates in Panel B are very similar to those in column 1.[21]

Absent income data by year and county (or household) from the 1970s to 1980s, these estimates on grain output are helpful to understand the lower bound of income change fellowing the reform. Nationwide, changes in procurement price of grain products by year had been universal during this period, which are accounted for by year fixed effects. Thus, the increase in individual income from selling grain products would be at most 2.6 percent. In addition, from a small sample of counties that report year-by-year data on production of both grain and cash crops, we find that the fraction of grain output only decreases slightly from 98 percent in 1974 to 93 percent in 1984.[22] Previous literature suggests that land reform led to a more dramatic increase in production of cash crops. However, because on average cash crop was still a very small proportion in agricultural production by 1984, we would expect a tight upper bound of income increase.

The estimated "first stage" effect of land reform on grain output has two implications for our empirical analysis below. First, it provides us with the guidance to assign the timing of exposure status at the individual level. If one was born two years after the reform, he/she was exposed to a positive income shock *in utero*, while those born one year after the reform or earlier were unexposed in the prenatal period. Second, because the productivity gain and therefore the income increase induced by land reform is small, we expect that average effect on individual academic performance generated by land reform is moderate.

[20] This is consistent with a consensus in the literature that China's agricultural performance slowed down after 1985 (Sicular, 1991; Lin, 1992).

[21] Our estimated effect of the HRS adoption on productivity gain is much smaller than the conventional estimate in Lin (1992). An important feature of our identification strategy is in its ability to control for time varying effects of county characteristics that are correlated with the timing of HRS adoption. Besides, two additional differences are noticeable. First, while Lin (1992) measures the HRS adoption by the fraction of production teams that had adopted the HRS by year for each province, we use the year HRS was first introduced for each county. Second, the outcome in Lin (1992) is the value of crop output including grain crops and cash crops, while we only use grain output. Taken at face value, our findings using a more careful research design and data at a more local level cast some doubt on the large effect of HRS adoption on agricultural productivity found in the previous literature.

[22] Only 36 counties are in this sample. Cash crops include cotton lint, oilseed, sesame, peanut, etc. We do not use this sample for estimation because of the small sample size.

2.4 Education system in China

Pre-college education in China includes six years in primary school, three years in middle school and three years in high school. Education in primary school and middle school has been mandatory since the introduction of a nine-year compulsory schooling law in 1986. In the final year of middle school, students take the local entrance exams to get access to high school. Individuals who fail the exams either move to vocational schools for 3-5 years of study, or stop going to school. In most provinces, high school students are divided into two tracks: science and humanity/social science. Math, Chinese and English are the core subjects for both tracks.[23] In addition, the science curricula include physics, chemistry and biology, while the humanity/social science curricula include history, geography and political education.

The procedure of admission to college is comprised of two stages. In the first stage, high school students take the national college entrance exams in their final year. These exams last for 12 hours over 3 days from June 7th to 9th. Math, Chinese and English are the three basic subjects. Depending on the track in high school, an additional subject that combines respective subjects in each track is required. A test taker's overall score is a weighted sum of scores from these four subjects. While the exams are standard for most provinces, there are a few exceptions. Beijing and Shanghai are allowed to design independent exams, and more provinces start to have their own exams in recent years. Within province, the difficulty of the exams also varies between science track and humanity/social track. For example, math exam is designed to be more difficult for students in the science track, while Chinese exam is relatively easier.

In the second stage, one's admission to college is determined by his/her total score on the national exams and his/her preference on colleges. Depending on the provincial policy, students fill application forms ranking their most preferred colleges, either before or after the national exams. Colleges set a fixed enrollment quota for each of the two tracks for each province. To be admitted by a particular college, a student has to pass the score threshold of that college for his/her track in the province he/she takes the exams, and has listed that college in his/her top choices. Thus, student essentially compete against each other within their track in their province.

On the supply side, there is a clear hierarchy among the 2,300 higher education institutions in China. At the bottom of the pyramid are three-year colleges that typically award associate degrees. Above them are four-year colleges offering programs in both academic and vocational courses and typically leading to bachelor degrees. Four-year colleges are classified into first- and second- tier, the former of which have better reputation and enroll students with higher scores in the college entrance exams. On the top of the pyramid are 39 elite colleges and 100 top colleges that most of the government research fundings spend on.[24] Only test takers at the higher end of the score distribution have access to these colleges that set the highest thresholds in admission.

[23] Other foreign languages are taught in a few areas, e.g. Russian, Korean, etc.

[24] People's Daily Online, "Over 10 billion yuan to be invested in *211 Project*". http://english.people.com.cn/90001/6381319.html.

3 Data

The primary microdata we use cover the universe of applicants in college entrance exams between 1999 and 2001. Applicants report their national identification number, from which one's county of *hukou* registration and date of birth are precisely obtained. Because one*'s hukou* status is registered before age 1, and internal migration was strictly restricted under the *Hukou* system from the 1970s to the mid 1980s, the county of *hukou* registration corresponds well to the *in utero* county for individuals born during this period. This is an important advantage of our research design. Migration for college entrance exam has been widely reported by media. Some parents send their children to high schools in less developed provinces where the score thresholds in college admission are set lower. In this setting, having a good measure of one's *in utero* county largely reduces concerns about possibly endogenous mobility.

For each year between 1999 and 2001, we match the microdata with the county-level data on the timing of land reform based on one's county of *hukou* registration. Although we observe the universe of college applicants in the microdata, the fraction of provinces that release the national identification number of applicants increases by year. In 1999, 13 out of 30 provinces release national ID number, 22 provinces in 2000 and 29 in 2001.[25] There are 472 counties that we can match in both the microdata in 1999 an the county-level data on reform timing, 719 counties in 2000 and 1044 in 2001. In total, we have 1068 counties matched in both data. The dotted line in Figure 1 shows that the distribution of the land reform timing in these matched counties is very close to that in all counties that have ever reported the reform timing. We also show in Appendix Table 1 that, in the 1990 and 2000 Census, these matched counties are very similar to unmatched counties in education level, including high school completion rate and fraction of population with some college. The only difference is that matched counties have less labor force in nonfarm employment, suggesting that the fraction of families that benefited from land reform might be slightly higher in our sample of counties.

Our analysis sample includes test takers born between 1978 and 1984, because the vast majority of test takers during 1999-2001 were at age 17-21 (98%). An individual's *in utero* exposure to land reform is jointly determined by one's year of birth and the timing of HRS adoption in his/her county of *hukou* registration. We measure *in utero* exposure to land reform as whether one was born two years after the reform or later, and use variation across cohorts and counties to isolate the effect of prenatal exposure on academic performance. In our sample, there are 43.8 percent of test takers who were exposed to land reform *in utero* (the HRS remains in place once adopted). They were born from five years before the reform to six years after the reform.

We consider two sets of measures on academic performance as the main outcomes. First, we focus on the total score and scores of the three core subjects: math, Chinese and English. Because the exams might vary across provinces and also between science track and humanity/social science

[25] Guangdong doesn't release national ID number in all years.

track within province, all test scores are standardized within province and track.[26] In our empirical analysis, we use the percentile of test score for each subject relative to the sample of test takers within province and track. Second, we examine a set of admission outcomes, including whether one was admitted by any college, by a 4-year university, by a first-tier university, by a top 100 university or by one of the 39 elite universities. Our analysis sample includes 1,918,244 complete records on total score and admission outcomes.[27]

Table 3 reports summary statistics of the analysis sample. While total scores of male and female students are very close, the variation among male students is larger. Male students perform slightly better in math exam, while female students perform better in Chinese and English exams. On admission outcomes, more than half of all test takers, 53 percent, are admitted by any college. There are 25.1 percent admitted by a four-year university and 10.6 percent admitted by a first-tier university. A small proportion of them, 6.7 percent, have access to a top 100 university, and only 2.4 percent go to an elite university. While the ratio of being admitted by any college is slightly higher among female test takers, male test takers perform better in terms of being admitted to higher ranked universities.

4 Empirical Framework

4.1 Identification strategy

To infer the causal relationship between *in utero* exposure to land reform and later-life academic performance, two major challenges exist. The first challenge is to overcome the identification problem: omitted variables which might be correlated with one's exposure to land reform. The second challenge concerns sample selection. The sample of test takers might be a selective one if land reform affects fertility, or one's chances of going to high school, or one's decision to apply for college and take the national exam, or all of them. Assuming for the moment that land reform has no impact on one's selection into the sample of test takers, we first consider the empirical approach to isolate the effect of exposure to land reform, and then discuss the sample selection problem in the next subsection.

Our identification strategy compares individuals who were *in utero* just before the first post-reform harvest and those just after. One's exposure status is jointly determined by one's year and county *in utero*. As suggested by our "first stage" estimate on grain output, the first cohort exposed to the positive income shock was born two years after the reform. Using a difference-in-difference approach, we estimate equation (1) as below:

[26] We use the official formula of standardization used by the Ministry of Education.

[27] Among these test takers, 99.5 percent have records on math and Chinese scores and 99.4 percent have records on English score.

$$Y_{ijte} = \alpha + \beta Reform_{jt} + \gamma_j + \delta_t + \eta_e + \phi_j * t + D'_{jt}\theta_t + D'_{jt-1}\lambda_{t-1} + \sum_{t=1979}^{1984}(X'_j * T_t)\rho_t + \varepsilon_{ijte} \quad (1)$$

where i denotes the individual, j the *in utero* county, t the year of birth and e the year of college entrance exams. Y_{ijte} denotes one's academic performance, as captured by test scores and admission outcomes. The variable of interest, $Reform_{jt}$, is equal to 1 if one was born two years after the land reform or later and 0 otherwise. A set of fixed effects, county fixed effects γ_j, year of birth effects δ_t and the exam year effects η_e, are controlled for. These county and time controls absorb the effects of time invariant county characteristics, national birth cohort effects, as well as any differences in the college entrance exams and admission process across provinces or over time. Moreover, we also include county specific linear trends, $\phi_j * t$, to account for county characteristics that change smoothly over time and that are correlated with the timing of land reform. Because the identifying assumption is that cohorts born prior to the reform give the right counterfactual, including county specific linear trends reduces concern about pre-existing trends that differ between counties that reformed earlier and those that reformed later.

Furthermore, we account for time varying effects of county characteristics that are found to drive the reform timing. The time-variant measures on drought in March and April of the current year are denoted by D'_{jt}, and drought of last year is denoted by D'_{jt-1}. The time-invariant determinants prior to 1978, including labor force density in 1977, famine intensity in 1959-61 and distance to Beijing, are denoted by X'_j. We interact these covariates with time fixed effects from 1979 to 1984, with 1978 omitted.

A specific concern arises if other post-Mao reforms during the same period also affect long-term human capital. Particularly, because the introduction of HRS came as part of the whole package of market-oriented reforms in the rural economy, the other two major reforms, namely, price reform and market reform (Lin, 1992), could also increase household income and improve later-life human capital. To confound the effects of land reform, other reforms should have followed the timing of land reform closely at the county level and have had differential impacts on individuals born just before and just after two years after land reform. However, the increases in procurement prices and in bonuses for above-quota productions were nationwide in 1979 (Sicular, 1991), and the reductions in the planning of agricultural production and in the restrictions in interregional trade were universal state interventions (Lin, 1992). Any effects of exposure to these two reforms are accounted for by controlling for year of birth effects δ_t.

Besides economic reforms, we also conducted a comprehensive reading on non-economic reform policies from the late 1970s to the mid 1980s. We are concerned that the One Child Policy introduced in 1979 might be a potential confounder because parents might invest more in the prenatal period under the fertility restriction. OCP rules varied over time from 1979 to the present. For our purpose, the threat comes from the initial enforcement of OCP by county between 1978 and 1985 that might overlap with the county-level rollout of land reform. We have compiled the most detailed data on the timing of OCP rollout by county. Results for the effect of land reform on academic performance is robust when exposure to OCP is controlled for.

Another concern is that prenatal exposure to land reform might affect one's age at the college entrance exams, which is equal to the exam year minus one's year of birth. For example, if individuals exposed to land reform *in utero* perform better in their pre-college education and therefore their chances of repetition are lower, they would take the exams earlier than those unexposed. In the presence of such effect, including the exam year effects, η_e, would yield a "bad control" problem in the identification (Angrist and Pischke, 2008). To directly address this concern, we estimate equation (1) by using one's age at the college entrance exams as the outcome and provide evidence that it is not affected by one's exposure to land reform.

Finally, besides raising rural household income, land reform might have changed other household conditions. The change most relevant to child human capital is the collapse of rural medical system following the decollectivization. The coverage by collective health care of the rural population went down from 90 percent in the 1970s to 5 percent in the mid 1980s (Hsiao, 1995). If less access to health care has long-term effect on human capital, we would expect that the effect is negative, which is the opposite to that of a positive income shock generated by land reform.

4.2 Discussion on sample selection

We observe test scores and admission outcomes for high school graduates who took the national exam. Selection into this sample might be endogenous to one's exposure to land reform for four main reasons. Fertility response to land reform is the first one. A common concern in the fetal origins literature is about the scarring effect (Almond and Currie, 2011b), implying negative selection of our sample induced by land reform that might reduce fetal mortality. Moreover, parents might desire more children after the reform. The shift from collective farming to household farming increased the value of manual labor and therefore household demand for labor. The allocation of land is based on family size, which could also provide parents incentive to have more children. Empirically, Schultz and Zeng (1999) find little effect of land reform on fertility in the 1980s. In Almond et al., (2012), we find that fertility responded positively to land reform. However, the magnitude of this effect is small, about 2%. The scale of the fertility response is sufficiently modest so as not to constitute an alternative "selection story".

Second, if exposure to land reform improves one's performance in the high school entrance exams, one is more likely to go to high school. Similarly, they are less likely to drop out if they perform better in high school. Thus, the sample of high school graduates would be positively selected. In the 2000 Census, there are 16 percent of individuals born in 1978-84 in our sample of counties who ever went to high school or were still in high school. Among them, only 2 percent dropped out in high school.

Third, high school graduates might choose not to apply for college upon graduation. If one's exposure to land reform affects one's decision to apply, the effect would also be positive because it is students who are not confident in performing well give up taking the exam. The national statistics show that 89 percent of high school graduates apply for college in 1999, 96 percent in 2000 and 93 percent in 2001.[28]

Finally, a few outstanding students are exempted from national exam and guaranteed college admission. If those exposed to land reform were the top students in high school, they are more likely to go to colleges, and better colleges, without taking exam. Among all applicants in our data, 96.8 percent of them took the national exam, which might be a negatively selected sample.

We can directly test three out of these four hypotheses. Using the 2000 Census, we can estimate the reduced form of exposure to land reform on whether one ever attended high school and high school dropout. Using the administrative data on test scores, we can test the selection from college applicants to test takers.

Because we do not observe high school graduates who do not apply to college, the only hypothesis that we can not directly test is about the selection from high school graduates to applicants. Instead, we use the 1990 and 2000 Census to obtain the size of high school graduates by county, year of birth and year of graduation, and match these counts with the counts of college applicants in the administrative data.[29] In each county-birth cohort-year of graduation cell, we test whether the probability of applying to college is affected by exposure to land reform.

5 Main results

5.1 Results

We report the estimated effects of *in utero* exposure to land reform on academic performance, as captured by test scores in percentile and admission outcomes, in Table 4. Exposure to land reform increases one's total score by 0.25 percentile, which is statistically significant at 10 percent level. The effect size is 0.5 percent compared to the sample mean, a very small magnitude. Similar effects are found on math and English scores, and the effect on math score is slightly larger (a 0.6 percent increase) and statistically significant at 5 percent level. While there is no strong evidence that land reform affects one's chances of going to college or a four-year college, test takers exposed to the reform are 2.8 percent more likely to be admitted by a first-tier college and 4.5 percent more likely by a top 100 college, the latter of which is statistically significant at 1 percent level. The larger effect on going to better colleges might suggest that exposure to a positive income shock disproportionately benefits individuals with better endowments *in utero*. Another possibility is that parents from higher socioeconomic background invest more prenatally, or in childhood, or

[28]Educational Statistics Yearbook of China.

[29]See Data Appendix on the construction of the cell size.

both, as responses to an exogenous increase in household income.[30]

In Table 5, we report the results by gender. On test scores, while estimates for female students are all statistically insignificant and small, male students benefit from exposure to land reform in their total score, math score and English score. Particularly, the estimated effect on math score for boys is more precise (statistically significant at the 1 percent level) and relatively larger in magnitude (a 0.8 percent increase). We also find little effect of land reform on admission outcomes for girls, except for a 3.5 percent increase in the probability of going to a top-100 college at the 10 percent significant level. For exposed boys, they are 4.5 percent more likely to be admitted by four-year colleges, and 5.5 percent more likely by top 100 colleges, which are larger effects compared to these in the full sample. Moreover, they are 7 percent more likely to go to one of the 39 elite colleges, the very top of the higher education pyramid, though the estimate is statistically significant at the 10 percent level. Overall, the positive effect of prenatal exposure to land reform on performance on college entrance exam largely loads on boys.

5.2 Interpretation of gender difference

It is not obvious whether the gender difference is purely biological or may be attributed to gender discrimination in parental investments. If male fetuses are more sensitive to the nutritional improvement induced by land reform, we would expect a lower fetal mortality rate for boys and thus a higher sex ratio after the reform. In Almond et al., (2012), we find that the sex of the first child is not affected by land reform, suggesting a minimal role of biology in explaining the gender difference. Alternatively, in Model Appendix, we formalize two ideas for understanding the role of parental behaviors in the context of a society with son preference. Following Almond and Currie (2011a), we consider that human capital is produced in a two-period childhood. The shock *in utero* –μ_g in the model – is the same for boys and girls. As suggested by Chen et al., (2011), less than 20 percent of all counties had introduced ultrasound by 1982, when more than 98 percent had started land reform. Prenatal investments are therefore unlikely to respond differentially to the shock for boys and for girls. When h is the human capital after childhood (e.g. math score), we show that son preference in postnatal childhood investments will lead to

$$\frac{\delta h^{(b)}}{\delta \mu_g} > \frac{\delta h^{(g)}}{\delta \mu_g},$$

where b, g refer to boys and girls, respectively.

We first assume that the level of postnatal investment is always higher for boys than for girls (subsection A of model appendix). For less than perfect substitutability between developmental periods, boys benefit more from the shock in terms of later-life human capital. The difference is generated mechanically by the gender discrimination in the investment level in childhood, without

[30] Unfortunately, parental information is not available in the administrative data on test scores.

assuming optimization. Intuitively, if postnatal investments are high for boys, then prenatal investments are a more binding constraint on human capital (to the extent that the substitutability in investments between prenatal and postnatal periods is imperfect), so land reform relieves the production constraint more for boys.

Second, we consider postnatal investments that are endogenous to the land reform shock (subsection B of model appendix). We build on Almond and Currie (2011a) but assume here that parents value perceived income of their child instead of child's human capital *per se*, and the perceived return to human capital is higher for boys than for girls.[31] The Chinese Longitudinal Healthy Longevity Survey in 2005 shows that the rural elderly rely more on sons both financially and emotionally.[32] We illustrate that, given a positive shock *in utero,* parents respond more in postnatal investments for sons than for daughters. Thus, the shock would have larger effect on improving academic performance for boys.

In sum, both interpretations are consistent with our reduced form evidence on the gender difference. A final note is on the effect for girls. Regardless of gender discrimination in parental investments, we might still expect a positive effect for girls through the biological channel. Estimates for girls have similar standard errors but smaller point estimates relative to boys. Although these estimates are imprecise, we can not reject certain positive effects for girls. Moreover, if the production function is Leontieff in periods and the postnatal investment for girls is zero, we would observe zero effect of prenatal exposure to land reform for girls.

6 Robustness

6.1 One Child Policy

One Child Policy was officially announced by the central government in January of 1979. Distinct from the birth control policies that had relied on press campaign and propoganda since the early 1970s, the 1979 policy introduced a new system of financial incentives for birth control for the first time. Under the initial policy, parents who sign the one-child pledge would get rewards, while those who give birth to three or more would suffer economic sanctions (Banister, 1987).[33] At the local level, the implementation of these rules was in a highly diverse manner (Scharping, 2003).

We have compiled new data on the year the county government issued the first policy document

[31] It would be equivalent to assume that parents value a unit of human capital accumulation by sons more than daughters in our framework. In the 1991 Survey on Women's Status in Contemporary China, 73 percent of rural women think that a son should have at least high school education, while 54 percent think so for a daughter.

[32] Among rural individuals aged over 60 who live with children, 92% of them live with sons. The financial transfer they get from sons is 1.7 times of that from daughters. When they are sick, 56% are taken care of by sons and daughter-in-laws, while only 9% by daughters and son-in-laws. Moreover, for 66% of the elderly, the first person they talk to if they have difficulties is son or daughter-in-law, while only 4% talk to daughter or son-in-law.

[33] The initial policy did not set a rule for the second child. The policy was tightened to allow only a few types of families to have the second child in 1982, and it was relaxed in 1984 to issue the second child permit to a broader range of families (Greenhalgh, 1986; Scharping, 2003).

to enforce rewards for the single child and penalties for above-quota births, also from county gazetters. In Figure 3, we plot the distribution of the timing of land reform and OCP from 856 counties that report the timing of both reforms. There are 56 percent of counties that enforced OCP in 1979, while 1 percent started as early as 1978 and the other 43 percent implemented between 1980 and 1985. Introduced during the same period, land reform and OCP show substantial difference in their timing at the county level, which allows us to separate the effect of land reform from that of OCP.

If the introduction of OCP induced an increase in prenatal investment of parents, we would expect that individuals *in utero* in the year when OCP came in were the oldest cohort affected. Thus, we assign treatment status to OCP as 1 for individuals born one year after OCP or later and 0 otherwise. In Table 6, we report the estimates for both exposure to land reform and exposure to OCP. Recall that the sample of counties is smaller than our full sample of counties used for Table 4 and 5. For estimates on test score in this sample, only the estimate for total score is statistically significant at the 10 percent level. Despite less precise estimates, we find that the estimate for exposure to land reform has little change after exposure to OCP is controlled for. Results for admission outcomes are similar to those in the full sample and are robust. Table 7 and Table 8 report results for boys and girls, respectively. Estimates for both boys and girls are very similar to those in Table 5, and again they are robust to the inclusion of exposure to OCP.

These findings suggest little parental responses in the prenatal period to the One Child Policy and that OCP does not confound our results for the effect of land reform.

6.2 Testing sample selection

This subsection reports results from testing possible concerns about sample selection discussed in Section 4.2. In Table 9, column 1 and 2 report results on testing selection into high school graduates. We use the 2000 Census because individuals born between 1978 and 1984 were at age 16-22 in 2000.[34] There is suggestive evidence that *in utero* exposure increases one's probability of going to high school by 0.5 percentage points, a 3 percent increase relative to the sample mean. However, the estimate is not statistically different from zero. Consistent with the expectations, exposure to land reform reduces high school dropout. The estimate is also very small and statistically insignificant.

Next, column 3 shows whether a high school graduate's decision to apply for college is affected by his/her exposure to land reform. Combining the 1990 Census, the 2000 Census and the administrative data on test scores, we conduct this test by county, birth year, and year of application. The estimate is extremely small, a decrease of 0.1 percentage points relative to the sample average of 39 percent. Again, it is imprecisely estimated and with the wrong sign.

[34]The 2000 Census is matched with the data on the land reform timing using one's county of residence in 2000. One caveat of using the 2000 Census is that migration rate has been rising since the relaxation of *Hukou* system in the 1990s. The migration rate is 8.4 percent in our sample.

We report the result for testing the selection from applicants to test takers in column 4. As expected, applicants exposed to land reform are less likely to take the national exam, meaning that they are more likely to be guaranteed admission. This is consistent with our finding on positive effect of exposure to land reform on academic performance. However, the estimate is economically very small and statistically insignificant.

To briefly summarize results from column 1 through 4, we find little evidence that prenatal exposure to land reform affects one's probability of being observed in the sample of test takers in the college entrance exams, reducing concerns about endogenous sample selection.

Lastly, recall the concern about one's age at exam and the potential "bad control" problem as discussed in Section 4.1. In column 5, we use the administrative data to examine the effect of exposure to land reform on one's age at the college entrance exams. Test takers exposed to land reform is 0.005 year (roughly 2 days) younger than those unexposed. The effect is economically small and statistically insignificant, suggesting that including year of exam effects would not yield to bias in our estimation.

7 Conclusion

TBA

Model Appendix

A. Mechanical difference

Following Almond and Currie (2011a), we consider a simple two-period childhood. Human capital h at the completion of childhood is produced with a CES technology:

$$h = A[\gamma(\bar{I}_1 + \mu_g)^{\phi} + (1-\gamma)I_2^{\phi}]^{\frac{1}{\phi}}, \tag{2}$$

where I_1 denotes investment in the first period of childhood *(in utero* in our context), and I_2 denotes investment in the second period (postnatal period). An exogenous shock to (predetermined) period 1 investment is denoted by μ_g. Because ultrasound technology had been largely unavailable, for simplicity, we assume that prenatal investment is fixed.

In the presence of son preference, we assume that the investment level in period 2 is higher for boys than for girls, that is, $I_2^{(b)} > I_2^{(g)}$, and parental investments in period 2 do not respond to the shock. The impacts of the shock on human capital for boys and for girls are:

$$\frac{\delta h^{(b)}}{\delta \mu_g} = A[...]^{(b)\frac{1}{\phi}-1}\gamma(\bar{I}_1 + \mu_g)^{\phi-1} \tag{3}$$

$$\frac{\delta h^{(g)}}{\delta \mu_g} = A[...]^{(g)\frac{1}{\phi}-1}\gamma(\bar{I}_1 + \mu_g)^{\phi-1} \tag{4}$$

The sign of the gender difference in human capital as response to the shock depends on ϕ. For $\phi < 1$, (3) > (4), the shock has a larger effect in promoting human capital for boys.

B. Optimization

Now we consider parental responses to the shock. Parents value their consumption C and perceived income of their child y_c:

$$U_p = U(C, y_c) = B[\theta C^{\varphi} + (1-\theta)y_c^{\varphi}]^{\frac{1}{\varphi}}, \tag{5}$$

where $y_c = mh$ for boys and $y_c = nh$ for girls. We assume that the perceived return to human capital is higher for boys than for girls, that is, $m > n$. We also assume that parents view child as consumption instead of investment, i.e. they do not borrow against. Their budget constraint is:

$$\bar{I}_1 + I_2 + C = \bar{y}_p. \tag{6}$$

Let's consider boys first. Absent discounting, the marginal utility from consuming equals the marginal utility from investing:

$$\frac{\delta U}{\delta C} = \frac{\delta U}{\delta y_c}\frac{\delta y_c}{\delta I_2^*} \tag{7}$$

$$\theta(\bar{y}_p - \bar{I}_1 - I_2^*)^{\varphi-1} = (1-\theta)(mh)^{\varphi-1} mA[...]^{\frac{1}{\phi}-1}(1-\gamma)I_2^{*\phi-1} \tag{8}$$

$$\theta(\bar{y}_p - \bar{I}_1 - I_2^*)^{\varphi-1} = m^{\varphi}(1-\theta)(1-\gamma)A^{\varphi}[...]^{\frac{\varphi-\phi}{\phi}} I_2^{*\phi-1} \tag{9}$$

$$G(I_{2,}^* \mu_g) = \theta(\bar{y}_p - \bar{I}_1 - I_2^*)^{\varphi-1} - m^{\varphi}(1-\theta)(1-\gamma)A^{\varphi}[...]^{\frac{\varphi-\phi}{\phi}} I_2^{*\phi-1} = 0 \tag{10}$$

$$\frac{\delta I_2^*}{\delta \mu_g} = -\frac{\frac{\delta G}{\delta \mu_g}}{\frac{\delta G}{\delta I_2^*}}$$

$$= \frac{m^{\varphi}(1-\theta)(1-\gamma)A^{\varphi}(\varphi-\phi)[...]^{\frac{\varphi-2\phi}{\phi}} I_2^{*\phi-1}\gamma(\bar{I}_1+\mu_g)^{\phi-1}}{\theta(1-\varphi)(\bar{y}_p - \bar{I}_1 - I_2^*)^{\varphi-2} + m^{\varphi}A^{\varphi}(1-\theta)(1-\gamma)[...]^{\frac{\varphi-\phi}{\phi}} I_2^{*\phi-2}\left[(1-\phi)-(\varphi-\phi)(1-\gamma)I_2^{*\phi}[...]^{-1}\right]}. \tag{11}$$

We define B to be $(1-\theta)(1-\gamma)A^{\varphi}(\varphi-\phi)[...]^{\frac{\varphi-2\phi}{\phi}} I_2^{*\phi-1}\gamma(\bar{I}_1+\mu_g)^{\phi-1}$,
C to be $\theta(1-\varphi)(y_p - \bar{I}_1 - I_2^*)^{\varphi-2} \geq 0$,
and D to be $A^{\varphi}(1-\theta)(1-\gamma)[...]^{\frac{\varphi-\phi}{\phi}} I_2^{*\phi-2}\left[(1-\phi)-(\varphi-\phi)(1-\gamma)I_2^{*\phi}[...]^{-1}\right]$.
Equation (11) is simplified as:

$$\frac{\delta I_2^*}{\delta \mu_g} = \frac{m^{\varphi}B}{C + m^{\varphi}D}. \tag{12}$$

For $\phi < \varphi$, $B > 0$. Define the component in bracket of D to be E: $(1-\phi)-(\varphi-\phi)(1-\gamma)I_2^{*\phi}[...]^{-1}$, and

$$(1-\phi) - \frac{(\varphi-\phi)(1-\gamma)I_2^{*\phi}}{[...]} = \frac{(1-\phi)\gamma(\bar{I}_1+\mu_g)^{\phi} + (1-\gamma)(1-\varphi)I_2^{*\phi}}{[...]} \tag{13}$$

For $\phi < 1$, $E > 0$, and thus $D > 0$. (12) is positive so that positive shocks in period 1 should be reinforced. Next, we compare (12) for boys and for girls:

$$\frac{\delta I_2^{(b)*}}{\delta \mu_g} - \frac{\delta I_2^{(g)*}}{\delta \mu_g} = \frac{(m^{\varphi} - n^{\varphi})BC}{(C + m^{\varphi}D)(C + n^{\varphi}D)}. \tag{14}$$

For $m > n$, (14) is positive, suggesting that parents invest more on boys than on girls as response to the shock. Finally, consider the impact of the shock on later-life human capital for boys and for girls separately:

$$\frac{\delta h^{(b)}}{\delta \mu_g} = A[...]^{(b)\frac{1}{\phi}-1}\gamma(\bar{I}_1 + \mu_g)^{\phi-1} + \frac{\delta h^{(b)}}{\delta I_2^{(b)*}}\frac{\delta I_2^{(b)*}}{\delta \mu_g} \tag{15}$$

$$\frac{\delta h^{(g)}}{\delta \mu_g} = A[...]^{(g)\frac{1}{\phi}-1}\gamma(\bar{I}_1 + \mu_g)^{\phi-1} + \frac{\delta h^{(g)}}{\delta I_2^{(g)*}}\frac{\delta I_2^{(g)*}}{\delta \mu_g} \tag{16}$$

For $\phi < 1$, $A[...]^{(b)\frac{1}{\phi}-1}\gamma(\bar{I}_1 + \mu_g)^{\phi-1} > A[...]^{(g)\frac{1}{\phi}-1}\gamma(\bar{I}_1 + \mu_g)^{\phi-1}$, and $\frac{\delta^2 h}{\delta I_2^2} > 0$. Therefore, (15) > (16).

Data Appendix

A. Precipitation data

We use the Global Surface Summary of Day Data produced by National Climate Data Center (NCDC). Throughout China, daily data on total precipitation amount (0.01 inches) are available from 225 weather stations from 1956 to 1964 and 536 stations from 1973 to 1984. In each year, we assign each county in the 1982 Census the precipitation data from the nearest weather station using longitude and latitude. Because the number of weather stations increases overtime, a county might be assigned different stations in different years, with relatively closer stations in more recent years.

To construct the measure of drought in March, for example, we first get the distribution of total precipitation in March from all years during 1956-1964 and 1973-1984 for each county. We then define drought in March as a binary variable that is equal to 1 if the monthly precipitation is below the bottom 20 percentile of the distribution for each county in each year and 0 otherwise. For drought in the whole growing season, we calculate the average monthly precipitation from March to September and use its distribution to define drought.

B. Construct size of high school graduates in census data

We first count the number of births born between 1978 and 1984 by county and year of birth in the 1990 Census. A strength of using the 1990 Census for cohort size is that internal migration was still under strict control in 1990. We use the 2000 Census to calculate the fraction of population that ever went to high school or still in high school by county and year of birth. By matching cells in the 1990 and 2000 Census, we get the counts of high school students in each county-birth year cell in the 1990 Census.

Next, to match with the administrative data at the county-birth year-year of application level, we also need to further narrow down the cell in the 1990 Census by county, birth year and year of graduation. However, we do not observe year of graduation in any census data. Instead, we assume the age distribution in college entrance exams during 1999-2001 is the same as that of high school graduates who were born in 1978-1984. We calculate the fraction of applicants at a certain age (from 17 to 21) from the administrative data by county, and assign the fraction to each cell in the 1990 Census by county and year of birth. We then use this fraction as the probability of graduating at a certain age in each county-birth year cell, and get the size of county-birth year-year of graduation cell in the 1990 Census.

Lastly, we take the count of applicants by county, birth year and year of application in the administrative data and match with the count of high school graduates by county, birth year and year of graduation in the 1990 Census. We divide the cell size of applicants by the cell size of high school graduates, and get the probability of applying for college in each cell.

A final note is that we use the sample of first-time applicants (70%) in the administrative data for this test. Because we do not observe whether one repeat the last year in high school in census data, we only observe the cell size of high school graduates assuming that they graduate once. What we are concerned about is whether the decision of a high school graduate to ever apply for college is endogenous to land reform, which can be test using the probability of first-time application.

References

[1] Adhvaryu, Achyuta and Anant Nyshadham. 2011. "Endowments and Investments within the Household: Evidence from Iodine Supplementation in Tanzania." manuscript, Yale University.

[2] Almond, Douglas, and Janet Currie. 2011a. "Human Capital Development Before Age 5." In *Handbook of Labor Economics*, Vol. 4b, 1315– 1486. Elsevier.

[3] Almond, Douglas, and Janet Currie. 2011b. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives*, 25(3): 153–72.

[4] Almond, Douglas, Lena Edlund, Hongbin Li, and Junsen Zhang. 2010. "Long-term Effects of the 1959-1961 China Famine: Mainland China and Hong Kong." *The Economic Consequences of Demographic Change in East Asia*. NBER-EASE Volume 19, Chapter 9, 321-350, University of Chicago Press.

[5] Almond, Douglas, Hongbin Li, and Shuang Zhang. 2012. "Income and Sex Selection: A Cautionary Tale of Land Reform and Sex Ratios in China." Working Paper.

[6] Angrist, Joshua and Jorn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

[7] Bai, Ying and James Kai-sing Kung. 2011. "Better Incentives or Stronger Buffer? Weather Shocks and Agricultural De-collectivization in China." Working Paper, Hong Kong University of Science and Technology.

[8] Banister, Judith. 1987. *China's Changing Population.* Stanford: Stanford University Press.

[9] Bhalotra, Sonia and Atheendar Venkataramani. 2011. "The Captain of the Men of Death and His Shadow: Long-Run Impacts of Early Life Pneumonia Exposure." Working Paper, University of Bristol.

[10] Chen, Yuyu, Hongbin Li and Lingsheng Meng. 2011. "Prenatal Sex Selection and Missing Girls in China: Evidence from the Diffusion of Diagnostic Ultrasound." Working Paper, Tsinghua University.

[11] Chinese Academy of Agricultural Sciences. 1984. *China's Administrative Division for Farming.* Agriculture Press. Beijing.

[12] Chung, Jae Ho, 2000. *Central Control and Local Discretion in China: Leadership and Implementation During Post-Mao Decollectivization.* Oxford University Press.

[13] Datar, Ashlesha, Rebecca Kilburn and David Loughran. 2010. "Endowments and Parental Investments in Infancy and Early Childhood." *Demography*, 47(1), 125–144.

[14] Field, Erica, Omar Robles, and Maximo Torero. 2009. "Iodine Deficiency and Schooling Attainment in Tanzania." *American Economic Journal: Applied Economics*, 1(4): 140–69.

[15] Greeghalgh, Susan. 1986. "Shifts in China's Population Policy, 1984-86: Views from the Central, Provincial, and Local Levels." *Population and Development Review*, 12(3): 491-515.

[16] Heckman, James J. 2007. "The Economics, Technology, and Neuroscience of Human Capability Formation." *PNAS: Proceedings of the National Academy of Sciences*, 104(33): 13250–55.

[17] Hsiao, William. 1995. "The Chinese Health Care System: Lessons for Other Nations." *Social Science & Medicine*, 41(8), 1047-1055.

[18] Kraemer, Sebastian. 2000. "The Fragile Male." *British Medical Journal*, 321: 1609-12.

[19] Lin, Justin Yifu. 1987. "The Household Responsibility System Reform in China: A Peasant's Institutional Choice." *American Journal of Agricultural Economics*, 69(2): 410-415.

[20] Lin, Justin Yifu. 1988. "The Household Responsibility System in China's Agricultural Reform: A Theoretical and Empirical Study." *Economic Development and Cultural Change*, 36(3), S199-S224.

[21] Lin, Justin Yifu. 1992, "Rural Reforms and Agricultural Growth in China." *American Economic Review*, 82 (1) :34-51.

[22] Maccini, Sharon and Dean Yang. 2009. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall." *American Economic Review*, 99(3), 1006-1026

[23] McMillan, John, John Walley and Lijing Zhu. 1989. "The Impact of China's Economic Reforms on Agricultural Productivity Growth." *Journal of Political Economy*, 97(4): 781-807.

[24] McMillan, John. 2002. *Reinventing the Bazaar: The Natural History of Markets*. W. W. Norton & Company.

[25] Meng, Xin, Nancy Qian and Pierre Yared. 2009. "The Institutional Causes of Famine in China 1959-61." NBER Working Paper 16361.

[26] Perkins, Dwight. 1988. "Reforming China's Economic System." *Journal of Economic Literature*, 26(2), 601-645.

[27] Scharping, Thomas. 2003. *Birth Control in China 1949-2000: Population Policy and Demographic Development.* London and New York: RoutledgeCurzon.

[28] Schultz, Paul and Yi Zeng. 1999. "The Impact of Institutional Reform from 1979 through 1987 on Fertility in Rural China." *China Economic Review,* 10(2): 141-160.

[29] Sicular, Terry. 1991. "China's Agricultural Policy During the Reform Period." in *China's Economic Dilemmas in the 1990s: The Problems of Reforms, Modernization and Interdependence*, vol. 1, Joint Economic Committee, Congress of the United States, US Government Printing Office, Washington.

[30] Van den Berg, Gerard J., Maarten Lindeboom, and France Portrait. 2006. "Economic Conditions Early in Life and Individual Mortality." *American Economic Review*, 96(1): 290–302.

[31] Wang, Fei-Ling. 2005. *Organizing through Division and Exclusion: China's Hukou System*. Stanford University Press, Stanford, California.

[32] World Bank. 2000. *China Overcoming Rural Poverty*. Washington DC.

[33] Xue, Susan. 2010. "New Local Gazetteers from China." *Collection Building*, 29(3), 110-118.

[34] Yang, Dali. 1996. *Calamity and Reform in China: State, Rural Society and Institutional Change Since the Great Leap Famine*. Stanford University Press.

[35] Zweig, David. 1987. "Context and Content in Policy Implementation: Household Contracts and Decollectivization, 1977-1983." in M. David Lampton, ed., *Policy Implementation in Post-Mao China*, Berkeley: University of California Press.

Figure 1: Land reform timing

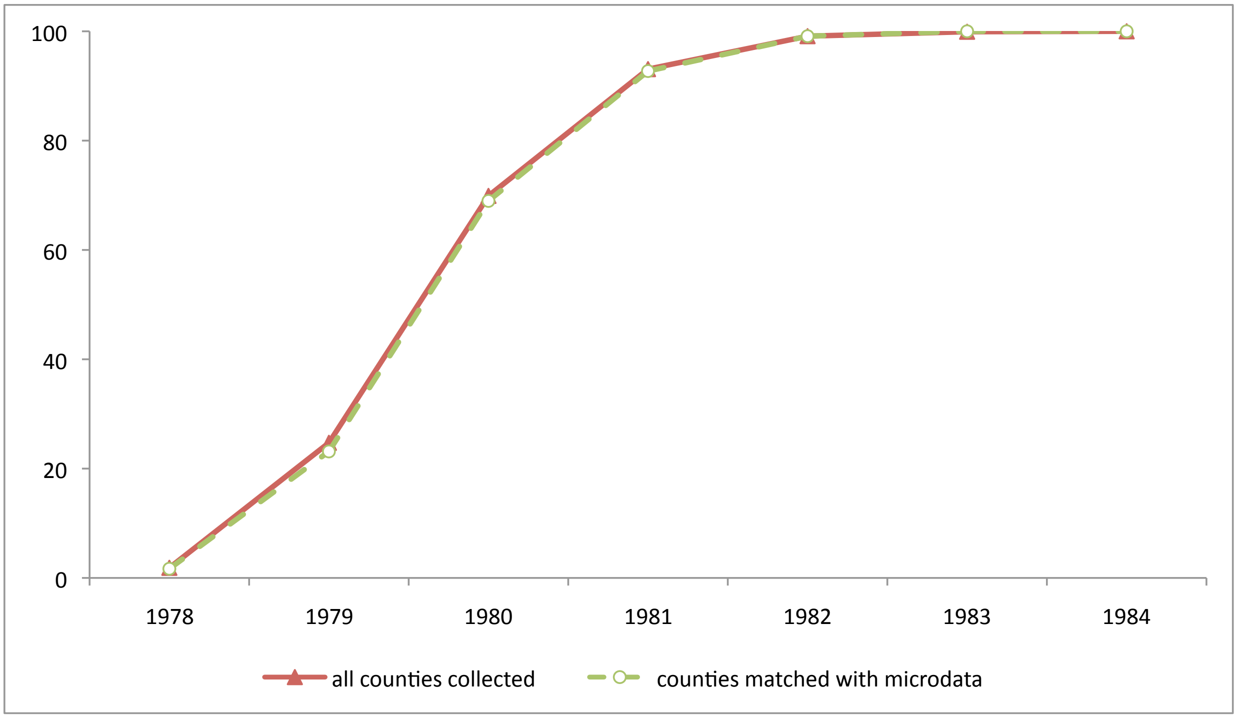

Figure 2: Land reform and grain output per capita

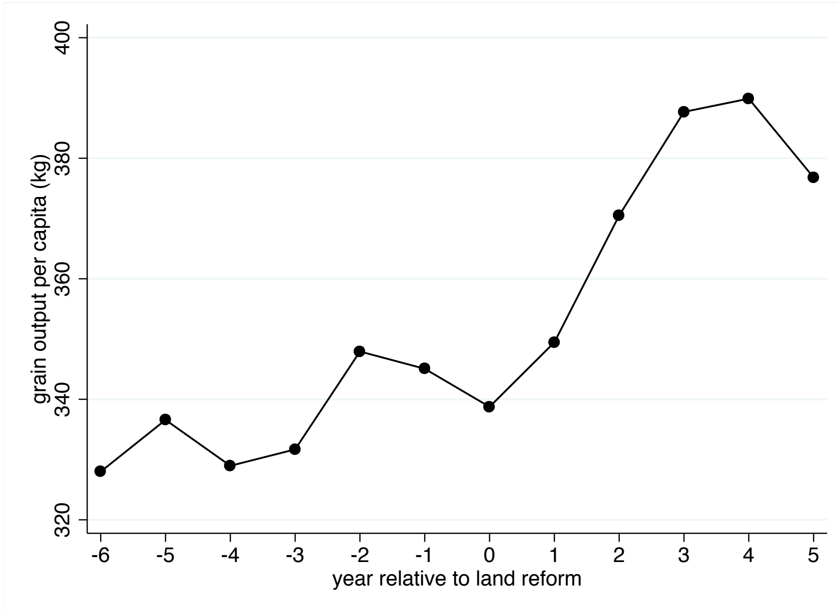

Figure 3: Timing of land reform *vs.* One Child Policy

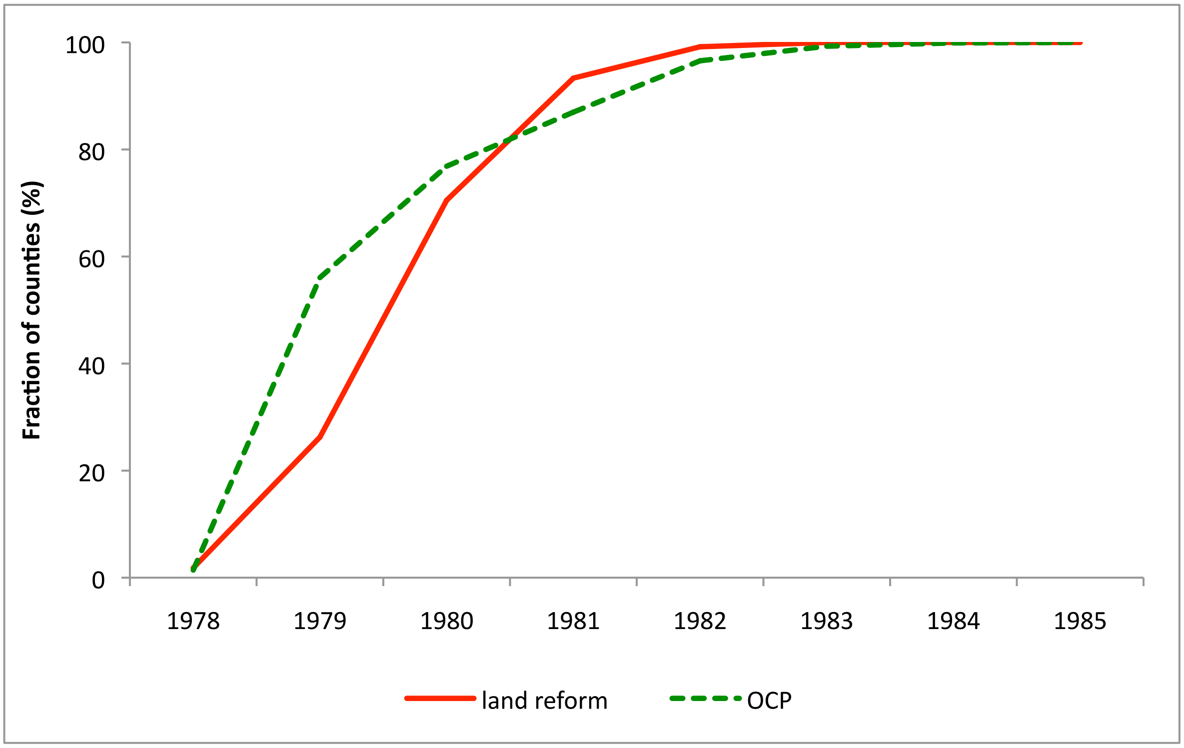

Table 1A: Drought (time-variant) and reform timing

	Dependent variable: first year of land reform (1978-1984)				
	(1) March-September	(2) March	(3) April	(4) May	(5) June
Drought in year t	-0.011	-0.021***	-0.037***	-0.006	-0.004
	[0.009]	[0.008]	[0.008]	[0.008]	[0.009]
Drought in year t-1	0.001	-0.026***	-0.027***	0.004	-0.009
	[0.008]	[0.007]	[0.008]	[0.008]	[0.008]
County FE	X	X	X	X	X
Year FE	X	X	X	X	X
County linear trend	X	X	X	X	X
Observations	7,306	7,306	7,306	7,306	7,306
R-squared	0.768	0.769	0.769	0.768	0.768

Notes: The dependent variable is the first year of land reform, which is equal to 1 at the year reform started, 0 prior to the reform, and missing value after the first year. Drought is a dummy variable which is equal to 1 if the average monthly precipitation is below bottom 20th percentile in the precipitation distribution during 1957-1984 and 0 otherwise (precipitation data between 1965-1972 are unavailable). We include two drought indicators, one in the current year and another the year before. In the first column we measure drought using monthly average precipitation in the whole period between March and September, while each of the other column headings presents the single month in which drought is measured. All regressions include county fixed effects, year effects and county linear trends. The sample includes 1194 counties and the time span is from 1975 to 1984. Robust standard errors are reported in brackets.

Table 1B: Time-invariant determinants of reform timing

	Dependent variable: first year of land reform (1978-1984)				
	Univariate			Multivariate	
		Obs	R-squared		
ln (grain output per capita 1976)	0.250**	481	0.011	0.399***	
	[0.121]			[0.127]	
ln (distance to province capital)	0.075**	1,201	0.003	0.004	-0.037
	[0.036]			[0.061]	[0.039]
ln (labor force density 1976)	-0.147***	1,117	0.044	-0.173***	-0.150***
	[0.022]			[0.046]	[0.028]
ln (famine intensity 1959-1961)	-0.494***	1,189	0.033	-0.289**	-0.342***
	[0.081]			[0.144]	[0.089]
ln (distance to beijing)	-0.074*	1,201	0.003	-0.123	-0.130***
	[0.038]			[0.078]	[0.041]
Observations				438	1,114
R-squared				0.095	0.070

Notes: The dependent variable is the first year of land reform, which varies from 1978 to 1984. For univariate analysis, each estimate is from a separate regression. Multivariate regression includes all dependent variables. Data on grain output per capita in 1976 are collected from county gazetteers, and only 438 counties report this information. Distance to beijing and distance to province capital city are in kilometers and are obtained from GIS map of 1982 Census. Labor force density in 1976 is calculated by population size aged 16-60 in 1976 divided by area. Using 1982 Census, we measure the 1959-61 famine intensity by the average cohort size born in 1953-1957 divided by the average cohort size born in 1959-1961. Robust standard errors are reported in brackets.

Table 2: Land reform and grain output per capita

	dependent variable: ln(grain output per capita)	
	(1)	(2)
	Panel A: diff-in-diff estimate	
1{1 year after reform or later}	0.032**	0.026*
	[0.015]	[0.015]
	Panel B: one estimate for each year after reform	
1{1 year after reform}	0.056***	0.053***
	[0.018]	[0.018]
1{2 years after reform}	0.062**	0.055*
	[0.031]	[0.031]
1{3 years after reform}	0.102**	0.098**
	[0.045]	[0.045]
1{4 years after reform}	0.160**	0.161***
	[0.063]	[0.062]
1{5 years after reform}	0.155*	0.162**
	[0.084]	[0.083]
1{6 years after reform}	0.132	0.168
	[0.127]	[0.124]
County FE	X	X
Year FE	X	X
County linear trends	X	X
Initial control*Year FE		X
Spring drought in t and t-1		X
Observations (panel A)	4,188	4,188
R-squared (panel A)	0.868	0.874
Observations (panel B)	4,188	4,188
R-squared (panel B)	0.869	0.875

Notes: This table reports estimates of reform on log grain output per capita. Panel A reports the difference-in-difference estimate. Panel B reports one estimate for each year after the reform within the difference-in-difference framework. Regression in column (1) includes county fixed effects, year effects and county linear time trends. Column (2) further includes determinants of reform timing interacted with time fixed effects and drought in March and Spring in year t and t-1. The analysis sample includes 382 counties that report grain output data in 1974-1984 and that are matched with the data of control variables. Robust standard errors are reported in brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Summary Statistics

	All sample		Female		Male	
	Mean	Obs	Mean	Obs	Mean	Obs
Total score	500.07	1918244	499.35	724471	500.52	1193773
	[94.72]		[91.28]		[96.74]	
Math score	501.78	1908119	498.13	720391	503.99	1187728
	[94.86]		[92.23]		[96.36]	
Chinese score	496.27	1909643	502.35	722327	492.57	1187316
	[97.97]		[95.92]		[99.01]	
English score	495.86	1906641	513.78	721309	484.96	1185332
	[95.81]		[90.80]		[97.12]	
Percentile of total score	49.79	1918244	49.61	724471	50.04	1193773
	[27.91]		[27.3]		[28.3]	
Percentile of math score	50.29	1908119	49.28	720391	50.99	1187728
	[28.08]		[27.5]		[28.4]	
Percentile of Chinese score	48.86	1909643	50.78	722327	47.89	1187316
	[28.52]		[28.3]		[28.6]	
Percentile of English score	48.77	1906641	54.29	721309	45.66	1185332
	[28.01]		[27]		[28.3]	
admitted by any college	0.530	1918244	0.555	724471	0.514	1193773
	[0.499]		[0.497]		[0.500]	
admitted by a 4-year university	0.251	1918244	0.240	724471	0.258	1193773
	[0.434]		[0.427]		[0.437]	
admitted by a first-tier university	0.106	1918244	0.098	724471	0.110	1193773
	[0.308]		[0.298]		[0.314]	
admitted by a top 100 university	0.067	1918244	0.057	724471	0.073	1193773
	[0.25]		[0.231]		[0.260]	
admitted by an elite university	0.024	1918244	0.018	724471	0.028	1193773
	[0.153]		[0.133]		[0.164]	
in utero exposure to land reform	0.438	1918244	0.451	724471	0.429	1193773
	[0.496]		[0.498]		[0.495]	

Notes: standard deviations are reported in brackets. The sample includes test takers born between 1978 and 1984 and who took the college entrance exams in 1999-2001.

Table 4: Effects of exposure to land reform on academic performance

	Test score in percentile				Admission outcomes				
	Total	Math	Chinese	English	Any college	Four-year college	First-tier college	Top 100 college	Elite college
1{born 2 years after reform or later}	0.251*	0.308**	0.114	0.249*	0.002	0.002	0.003**	0.003***	0.001
	[0.140]	[0.130]	[0.132]	[0.146]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]
County FE	X	X	X	X	X	X	X	X	X
YOB FE	X	X	X	X	X	X	X	X	X
Year of exam FE	X	X	X	X	X	X	X	X	X
County linear trends	X	X	X	X	X	X	X	X	X
Initial control*YOB FE	X	X	X	X	X	X	X	X	X
Spring drought in t and t-1	X	X	X	X	X	X	X	X	X
effect size relative to mean	0.005	0.006	0.002	0.005	0.004	0.008	0.028	0.045	0.042
dependent variable mean	49.79	50.29	48.86	48.77	0.53	0.251	0.106	0.067	0.024
Observations	1,918,244	1,908,119	1,909,643	1,906,641	1,918,244	1,918,244	1,918,244	1,918,244	1,918,244
R-squared	0.056	0.049	0.045	0.060	0.074	0.053	0.046	0.029	0.018

Notes: This table reports estimated effects of prenatal exposure to land reform on test scores and admission outcomes. Each estimate is from a separate regression on one outcome. The sample include test takers born in 1978 and 1984 and who took the college extrance exams in 1999-2001. All regressions include county fixed effects, year of birth effects, year of exam effects, county linear cohort trends, initial controls interacted with year of birth effects and drought in March and Spring in the current year and one year before. We also report the effect size of land reform on each outcome relative to the dependent variable mean. Robust standard errors clustered at the county level are reported in brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Effects of exposure to land reform on academic performance, by gender

	Test score in percentile				Admission outcomes				
	Total	Math	Chinese	English	Any college	Four-year college	First-tier college	Top 100 college	Elite college
Panel A: Female									
1{born 2 years after reform or later}	0.131	0.160	-0.016	0.209	0.001	0.001	0.001	0.002*	0.0003
	[0.178]	[0.178]	[0.179]	[0.185]	[0.003]	[0.003]	[0.002]	[0.001]	[0.001]
effect size relative to mean	0.003	0.003	0.0003	0.004	0.004	0.004	0.010	0.035	0.017
dependent variable mean	49.61	48.28	50.78	54.29	0.555	0.240	0.098	0.057	0.018
Panel B: Male									
1{born 2 years after reform or later}	0.348**	0.408***	0.203	0.321**	0.003	0.003	0.005***	0.004***	0.002*
	[0.157]	[0.148]	[0.145]	[0.159]	[0.003]	[0.002]	[0.002]	[0.001]	[0.001]
effect size relative to mean	0.007	0.008	0.004	0.007	0.006	0.012	0.045	0.055	0.071
dependent variable mean	50.04	51.00	47.89	45.66	0.514	0.258	0.110	0.073	0.028
County FE	X	X	X	X	X	X	X	X	X
YOB FE	X	X	X	X	X	X	X	X	X
Year of exam FE	X	X	X	X	X	X	X	X	X
County linear trends	X	X	X	X	X	X	X	X	X
Initial control*YOB FE	X	X	X	X	X	X	X	X	X
Spring drought in t and t-1	X	X	X	X	X	X	X	X	X
Observations (female)	724,471	720,391	722,327	721,309	724,471	724,471	724,471	724,471	724,471
R-squared (female)	0.066	0.060	0.058	0.068	0.089	0.064	0.056	0.034	0.019
Observations (male)	1,193,773	1,187,728	1,187,316	1,185,332	1,193,773	1,193,773	1,193,773	1,193,773	1,193,773
R-squared (male)	0.055	0.048	0.046	0.062	0.069	0.051	0.045	0.030	0.021

Notes: This table reports estimated effects of prenatal exposure to land reform on test scores and admission outcomes. Panel A reports estimates for female students and Panel B for male students. Each estimate is from a separate regression on one outcome. The sample include test takers born in 1978 and 1984 and who took the college extrance exams in 1999-2001. All regressions include county fixed effects, year of birth effects, year of exam effects, county linear cohort trends, initial controls interacted with year of birth effects and drought in March and Spring in the current year and one year before. We also report the effect size of land reform on each outcome relative to the dependent variable mean. Robust standard errors clustered at the county level are reported in brackets.
* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Exposure to One Child Policy controlled for

	Test score in percentile							
	Total score		Math		Chinese		English	
1{2 years after land reform}	0.264*	0.264*	0.371	0.377	0.143	0.143	0.250	0.250
	[0.158]	[0.158]	[1.853]	[2.360]	[0.151]	[0.150]	[0.162]	[0.162]
1{1 year after OCP}		0.036		-0.054		0.137		0.038
		[0.172]		[0.201]		[0.164]		[0.181]
Observations	1,504,260	1,504,260	1,495,855	1,495,855	1,496,886	1,496,886	1,495,004	1,495,004
R-squared	0.056	0.056	0.048	0.048	0.045	0.045	0.060	0.060

	Admission outcomes							
	Any college		First-tier college		Top 100 college		Elite college	
1{2 years after land reform}	0.001	0.001	0.003*	0.003*	0.004***	0.004***	0.001**	0.001**
	[0.003]	[0.003]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]
1{1 year after OCP}		0.001		-0.001		-0.001		-0.000
		[0.003]		[0.002]		[0.001]		[0.001]
Observations	1,504,260	1,504,260	1,504,260	1,504,260	1,504,260	1,504,260	1,504,260	1,504,260
R-squared	0.075	0.075	0.046	0.046	0.029	0.029	0.017	0.017

Notes: Robust standard errors clustered at the county level are reported in brackets. This table reports estimated effects of prenatal exposure to land reform and One Child Policy on test scores and admission outcomes. The sample include test takers born in 1978-1984 and who took the college entrance exams in 1999-2001. All regressions include county fixed effects, year of birth effects, year of exam effects and county linear cohort trends.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Exposure to One Child Policy controlled for, for boys

	Test score in percentile							
	Total		Math		Chinese		English	
1{2 years after land reform}	0.361**	0.361**	0.444***	0.445***	0.197	0.197	0.339*	0.339*
	[0.176]	[0.176]	[0.168]	[0.168]	[0.163]	[0.163]	[0.177]	[0.177]
1{1 year after OCP}		0.001		-0.09		-0.054		0.072
		[0.179]		[0.172]		[0.168]		[0.185]
Observations	936276	936276	931263	931263	930760	930760	929347	929347
R-squared	0.056	0.056	0.048	0.048	0.045	0.045	0.062	0.062
	Admission outcomes							
	Any college		First-tier college		Top 100 college		Elite college	
1{2 years after land reform}	0.003	0.003	0.005***	0.005***	0.005***	0.005***	0.002**	0.002**
	[0.003]	[0.003]	[0.002]	[0.002]	[0.002]	[0.002]	[0.001]	[0.001]
1{1 year after OCP}		0.001		-0.002		-0.001		0
		[0.003]		[0.002]		[0.001]		[0.001]
Observations	936276	936276	936276	936276	936276	936276	936276	936276
R-squared	0.07	0.07	0.045	0.045	0.031	0.031	0.021	0.021

Notes: Robust standard errors clustered at the county level are reported in brackets. This table reports estimated effects of prenatal exposure to land reform and One Child Policy on test scores and admission outcomes for boys. The sample include test takers born in 1978-1984 and who took the college entrance exams in 1999-2001. All regressions include county fixed effects, year of birth effects, year of exam effects and county linear cohort trends.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: Exposure to One Child Policy controlled for, for girls

	Test score in percentile							
	Total score		Math		Chinese		English	
1{2 years after land reform}	0.138	0.138	0.135	0.135	0.089	0.091	0.167	0.167
	[0.204]	[0.204]	[0.202]	[0.202]	[0.205]	[0.204]	[0.207]	[0.207]
1{1 year after OCP}		0.038		-0.041		0.465*		0.003
		[0.250]		[0.253]		[0.249]		[0.254]
Observations	567,984	567,984	564,592	564,592	566,126	566,126	565,657	565,657
R-squared	0.066	0.066	0.060	0.060	0.058	0.058	0.067	0.067
	Admission outcomes							
	Any college		First-tier college		Top 100 college		Elite college	
1{2 years after land reform}	-0.000	-0.000	-0.000	-0.000	0.002	0.002	0.001	0.001
	[0.004]	[0.004]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]
1{1 year after OCP}		0.001		-0.001		-0.001		-0.000
		[0.004]		[0.002]		[0.002]		[0.001]
Observations	567,984	567,984	567,984	567,984	567,984	567,984	567,984	567,984
R-squared	0.088	0.088	0.055	0.055	0.034	0.033	0.018	0.018

Notes: Robust standard errors clustered at the county level are reported in brackets. This table reports estimated effects of prenatal exposure to land reform and One Child Policy on test scores and admission outcomes for girls. The sample include test takers born in 1978-1984 and who took the college entrance exams in 1999-2001. All regressions include county fixed effects, year of birth effects, year of exam effects and county linear cohort trends.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 9: Test sample selection at the individual level

	(1) high school=1	(2) high school dropout=1	(3) college applicant=1	(4) test taker=1	(5) age at the exams
1{born 2 years after reform or later}	0.005	-0.0003	-0.001	-0.0017	-0.005
	[0.003]	[0.003]	[0.009]	[0.0013]	[0.009]
County FE	X	X	X	X	X
YOB FE	X	X	X	X	X
County linear trends	X	X	X	X	X
Initial control*YOB FE	X	X	X	X	X
Spring drought in t and t-1	X	X	X	X	X
Year of exam FE			X	X	
dependent variable mean	0.15	0.02	0.39	0.968	19.09
Observations	422,076	59,607	6,832	1,980,665	1,918,244
R-squared	0.074	0.062	0.784	0.052	0.775

Notes: Column (1) reports the estimate on whether one ever went to high school or still in high school using the 2000 Census. Column (2) report the estimate on high school dropout in the sample of individuals who went to high school using the 2000 Census. Column (3) report the estimate on the fraction of college applicants in high school students at the county-birth year-year of application level. Column (4) reports the estimate on whether an applicant had taken the national exam using the administrative data on test scores. Column (5) reports the estimate on one's age at college entrance exam using the administrative data on test scores.The sample includes individuals born in 1978-1984 in counties matched with the county-level data on the reform timing as well as all county-level controls. Regressions in column (1)-(2) and (5) include county fixed effects, year of birth effects, county linear cohort trends, initial controls interacted with year of birth effects and droughts in March and Spring in the current year and one year before. Column (3) and (4) further control for year of exam effects. Robust standard errors clustered at the county level are reported in brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix Table 1: Test sample selection at the county level

	1990 Census			2000 Census		
	unmatched	matched	difference	unmatched	matched	difference
high school completion	0.076	0.072	-0.004	0.115	0.116	0.0008
	[0.064]	[0.06]	[0.003]	[0.06]	[0.05]	[0.002]
some college	0.006	0.006	-0.0002	0.012	0.011	-0.0003
	[0.012]	[0.018]	[0.007]	[0.009]	[0.009]	[0.0004]
nonfarm employment	0.19	0.15	-0.04***	0.36	0.33	-0.03***
	[0.22]	[0.19]	[0.01]	[0.17]	[0.15]	[0.007]
observations	842	1034		869	1032	

Notes: the full sample includes counties defined by administrative rule: the two-digit county code is between 21 and 80. Standard errors are reported in brackets.