



Does the one-child policy improve children's human capital in urban China? A regression discontinuity design



Xuezheng Qin^{a,1}, Castiel Chen Zhuang^{b,*}, Rudai Yang^a

^a School of Economics, Peking University, Beijing 100871, China

^b Department of Economics, University of Washington, Seattle, WA 98195, United States

ARTICLE INFO

Article history:

Available online 19 September 2016

JEL classification:

J13

D13

I24

Keywords:

One-child policy

Quantity-quality trade-off

Regression discontinuity design

Human capital

ABSTRACT

Qin, Xuezheng, Zhuang, Castiel Chen, and Yang, Rudai—Does the one-child policy improve children's human capital in urban China? A regression discontinuity design

This paper is the first to examine the causal relationship between China's One-Child Policy (OCP) and the long-term accumulation of human capital by a regression discontinuity design. Based on the 2005 China Inter-census Survey data, we observe a strong policy shock upon the probability of being a single child among people born around the starting point of OCP, which in turn significantly increases their educational attainment in adulthood. The results strongly suggest that there exists a quantity-quality trade-off as renewed by Becker, and the trade-off is more pronounced in the economically less developed regions and among families with the same-sex children. The results shed light on the re-understanding of China's family planning initiative as well as the application of regression discontinuity designs. *Journal of Comparative Economics* 45 (2017) 287–303. School of Economics, Peking University, Beijing 100871, China; Department of Economics, University of Washington, Seattle, WA 98195, United States.

© 2016 Association for Comparative Economic Studies. Published by Elsevier Inc. All rights reserved.

1. Introduction

China's fast increasing aging population and old-age dependency ratio present a major demographic challenge to the nation's sustainable economic development and put its decades-long population control policy under scrutiny. In particular, the relaxation of the One-Child Policy (OCP), which requires the majority of parents to have only one child, has become an issue of heated debate in recent years. Many studies argue that the continuation of this rigorous policy will exacerbate the problem of shrinking labor force and thus have a negative impact on China's economic development (Ma, 2010; Zhou and Yin, 2011), while the opposite point of view suggests that the sudden relaxation of this family planning initiative would hamper the long-term economic growth with escalating population burden (Liu and Lu, 2008; Huang and Sun, 2013). In theory, the controversy between the two opposing views stems from the fundamental query of whether there exists a trade-off between population size and the average level of human capital accumulation. If the increase in fertility is accompanied by a substantial reduction in the per-capita human capital stock, then the relaxation of the birth control policy can hamper China's long-term innovation capability and its economic growth; however, if the trade-off is modest or non-existent, then

* Corresponding author.

E-mail addresses: qin.econpku@gmail.com (X. Qin), zogcee@gmail.com (C.C. Zhuang), rdyang@pku.edu.cn (R. Yang).

¹ Fax: +86 1062754237.

the ease of population restriction, along with the improvement of technology, may not be disadvantageous to China's long-term economic development.

The query is commonly known as the quantity-quality (Q-Q) trade-off, which has been well discussed in the human capital literature since the seminal works of [Becker \(1960\)](#); [Becker and Lewis \(1973\)](#); [Willis \(1973\)](#) and [Becker and Tomes \(1976\)](#). According to the classic Q-Q trade-off theory, households make independent decisions on the number (quantity) of children and the human capital (quality) investment in each of their children, which are negatively correlated under the plausible presumptions that parents are impartial to any of the children, have limited resources to invest in them (binding budget constraints) and have difficulties in borrowing money externally (credit market failures). Under such presumptions, the negative relationship between quantity and quality can be seen as a result of “price effect” and “income effect”: on the one hand, with the improvement of economic conditions (e.g., increasing provision of public education), the shadow price¹ of the quality of children will decrease holding their quantity constant, causing the quality to rise; on the other hand, since the income elasticity of quality is arguably larger than that of quantity,² when the household budget constraints are relaxed over time (e.g., through a rise of wage rates), the average quality of children will increase while their quantity will initially remain unchanged and then decline eventually due to the higher shadow price of quantity (raising an additional child becomes more “expensive”). In the empirical studies, the quality of children or their level of human capital is usually measured by educational attainment or health status. Most empirical tests find negative effects of family size on the quality ([Glick et al., 2005](#); [Rosenzweig and Zhang, 2009](#); [Hatton and Martin, 2010](#); [Liu, 2014](#)), whereas others find no effects ([Black et al., 2005](#); [Park and Chung, 2012](#)) or even positive effects ([Qian, 2009](#); [Lordan and Frijters, 2013](#)).

The discrepancy between the theoretical prediction and the empirical findings may be due to the violation of the important presumptions. For example, based on the census data in Norway, [Black et al. \(2005\)](#) show that the effect of family size on children's education becomes negligible when the children's birth order is controlled. [Qian \(2009\)](#) uses the China Health and Nutrition Survey (CHNS) data to find that the school enrollment rate of the first-born child tends to increase with family size, and the effect is larger within households where children are of the same sex. These findings may suggest that parents treat their children unequally, e.g., they may set a higher standard of educational attainment for the first-born child or pay more attention to the education of sons than daughters. [Park and Chung \(2012\)](#) use a sample from the Matlab Health and Socioeconomic Survey (MHSS) in Bangladesh to find that the Q-Q trade-off is only significant for the first- and second-born children in a family. Based on a sample from the Young Lives Project (YLP) in Peru, [Lordan and Frijters \(2013\)](#) find that the relationship between family size and health outcome (height) is negative only when the pregnancy is unplanned whereas the relationship turns positive for the planned child-births. These findings further suggest that the conditions for the Q-Q trade-off can be more complicated than presumed, especially considering the endogeneity (self-selection) of parental fertility choices.

The empirical studies aiming to test the Q-Q trade-off hypothesis are constantly challenged by the endogeneity problems, which may arise from two sources: on the one hand, the quantity and quality of children are often affected simultaneously by some unobserved factors (e.g. parents' socioeconomic status or their gender preferences for children); on the other hand, reverse causality may also exist – the increase of offspring's quality can encourage parents to raise more children in return. In order to control for such endogeneity, instrumental variables (IVs) are frequently used. The commonly used IVs in the related literature generally fall into two categories – the birth of twins ([Glick et al., 2005](#); [Li et al., 2008](#); [Angrist et al., 2010](#); [Hatton and Martin, 2010](#)) and the implementation of birth control policies ([Qian, 2009](#); [Park and Chung, 2012](#); [Cameron et al., 2013](#)), both of which are correlated to family size but arguably exogenous to the fertility decisions. Since population control policies are implemented with various intensities (generally mild) among different countries, researchers often choose the birth of twins as IV. However, this IV also has its pitfalls: first, although the birth of twins will accidentally increase the family size, such incidents can be jointly determined with the quality of children by the unobserved genetic factors; second, parental preferences for later-born children may also change after the birth of twins, e.g., the short-term economic pressures of having an unplanned twin can change the parents' long-term plans of having more children and also their attitudes towards the human capital investment in the twins. In summary, the IV method is mainly limited by its strong requirement of “exclusion restrictions”, i.e., the IVs must be uncorrelated to the unexplained portion of the outcome variable (the level of human capital in our case).

Compared to the IV-based tests on the Q-Q trade-off hypothesis, there is a paucity of empirical studies on this topic relying specifically on regression discontinuity (RD) designs. The basic idea behind the RD designs is to find an observable discontinuity in a treatment variable (the conditional probability of having a single child in our case) and an outcome variable (children's educational attainment), both of which relate to an assignment variable (children's date of birth) and jump discontinuously at a cutoff point of the assignment variable. Near the cutoff point, the treatment can be seen as if it

¹ The shadow price of quality is the cost of a unit increase in quality holding the quantity constant; similarly, the shadow price of quantity is the cost of an additional child holding the quality constant ([Becker and Lewis, 1973](#)).

² Because the “payback period” and the opportunity cost of investing in an existing child is smaller than that of nourishing a new child, it is more cost-effective for the parents to “buy” child quality than to “buy” child quantity with a marginal increase in income, which in turn gives rise to the higher income elasticity of child quality.

Table 1.
Historical evolution of family planning in China after 1949.

Stages	Years	Contents
1	1949–1961	The concept of family planning was put forward, which triggered a series of public debates and drew nationwide attention in China.
2	1962–1978	The related birth control policies and family planning agencies were gradually formed after twists and turns.
3	1979–1991	Family planning became one of the basic national policies in China, when One-Child Policy (OCP) was put in place in late 1979.
4	1992–2000	With the stabilization and standardization of OCP, population growth and its burden were well managed.
5	2001–2012	The fertility rate was successfully maintained at a relatively low level, whereas the central government tried to comprehensively manage the emerging demographic problems caused by the strict family planning.
6	2013–present	Under the pressure of aging population, the central government relaxed its family planning nationwide by proposing a “single parent, double children” policy, which later turned into a two-child policy in 2015.

Notes: The classification of the stages is based on Yang (2003) and Xu (2010) according to the contents of the related policy documents, which to some extent reflects the so-called “documental politics” in China.

is randomly assigned, and under some plausible assumptions³ the change of the outcome can be attributed to the change of the treatment – thus their causal relationship can be identified in the vicinity of the cutoff point. Compared to the “exclusion restrictions” imposed by the IV method, the RD approach relaxes this strong assumption and replaces it with a more justifiable one that individuals have imprecise control over the assignment (i.e., individuals cannot precisely decide whether they are assigned to the treatment group or the reference group).

To the best of our knowledge, this paper is the first to test the Q-Q trade-off hypothesis using an RD design. The strong tightening of the birth control policies in China in late 1979 – the launch of OCP – presents a good natural experiment opportunity for us to explore the relationship among the jumps in several social indicators (e.g., population fertility and education). The “internal validity” of RD (i.e., having a reliable causal interpretation for the observations around the cutoff point) is considered to be superior to the other quasi-experimental approaches, even though its “external validity” (i.e., the applicability of the causal inference to the overall population) hinges on the strong assumptions that justify the extrapolation to localities outside the RD bandwidth. To achieve the sampling representativeness given the limitation of such “external validity”, our study is based on a large sample drawn from the 2005 Inter-census 1% Population Survey of China. Our findings not only speak well to the classic Q-Q trade-off theory by reassessing the largest-scale population control policies in the world’s modern history, but also facilitate a better understanding on the microeconomic mechanisms of human capital investments that is in turn implicational for future population policies in China and other developing countries.

The paper proceeds as follows: Section 2 briefly reviews the historical context of China’s family planning policies after 1949; Section 3 introduces the empirical strategy; Section 4 then describes the data and sample; Section 5 presents the estimation results, validity tests and robustness checks (including a falsification test); Section 6 carries out the subsample analyses in accordance to the Q-Q theoretical presumptions; Section 7 concludes the paper with policy recommendations and a brief discussion of the re-understanding of OCP.

2. Family planning and the one-child policy

2.1. Historical background

As the most populous country in the world, China embarked on a path of population birth control after its establishment in 1949. Since then, the national family planning initiative has gone through six major stages as summarized in Table 1.

In the first two stages, the family planning in China was comparatively mild. It was then strengthened in the third and fourth stages, and relaxed again in the last two stages. At the turning points of policy changes, the single-child ratio (defined as the ratio of children with no siblings to the total numbers of children born at the same period of time) may jump upwards or downwards, enabling an RD design if the discontinuous change in the variable are statistically significant (i.e., if the exogenous policy shock on the population fertility rate is strong). Compared to the relatively lax implementation of birth control policies in other countries,⁴ the OCP launched in late 1979 in China is strong and unprecedented in terms of the width and depth of its enforcement, which may lead to a sufficient proportion of compliers⁵ among the population for our estimation.

³ These assumptions include (1) the continuity of the outcome variable as a function of the assignment variable in absence of the treatment, (2) individuals are not sorted across the cutoff in response to the treatment, (3) the other baseline covariates are continuous at the cutoff. These assumptions will be tested diagnostically in Section 5.2.

⁴ The relatively lax implementation of birth control policies in most countries may also explain why the RD method was hardly used in identifying the policy effects. The IV approach, in comparison, can be applied in absence of discontinuous jumps in the fertility rates and thus is more commonly used in the related studies.

⁵ In our context, the compliers are the single children whose parents would have been able (both physiologically and psychologically) to raise more than one child if OCP was not implemented; they may also be loosely referred to as the compliant families hereinafter.

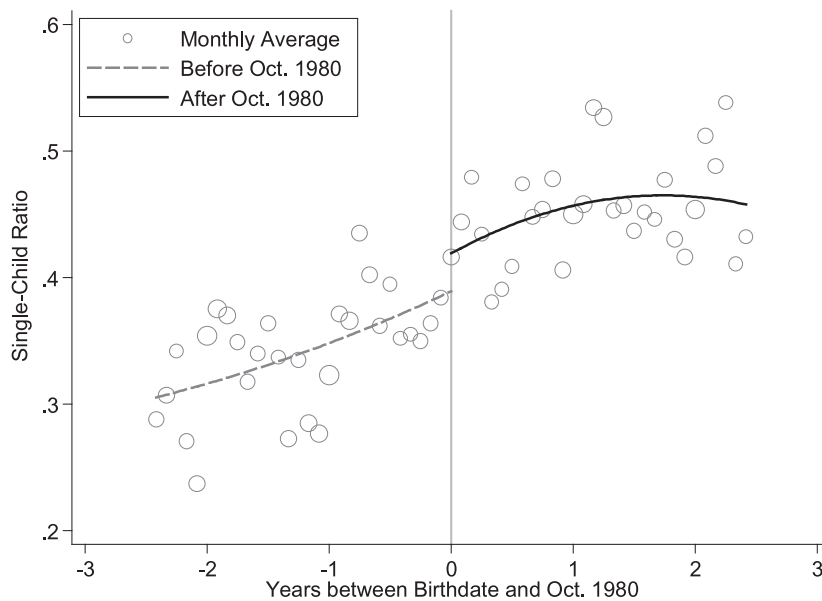


Fig. 1. Discontinuity in the single-child ratio (treatment variable) around Oct. 1980. *Notes:* (i) Data source: 2005 China's Inter-census Survey. (ii) The hollow circle represents the average single-child ratio in the corresponding month, and its size corresponds to the sample size; (iii) the solid line represents the quadratic predictions of the relationship between the single-child ratio and the adjusted month after October 1980, and the dashed line represents the relationship before October 1980.

2.2. Is there a policy cutoff point for RD designs?

OCP requires that each urban Han⁶ couple can only have up to one child (Zhang and Spencer, 1992; Li, 1995; Yang, 2003; Xu, 2010). The policy was implemented with “carrots” and “sticks”: on the one hand, the one-child certificates that offer a variety of benefits for the obedient couples started to be widely issued in November 1979; on the other hand, fines for the defiant couples also began to be imposed nationwide in January 1980.⁷ Thus, we assume that OCP was officially implemented during the late 1979 to the early 1980 period. According to Guo (2003), in our study period, OCP was enforced nationwide across more than half of the urban areas in China. Therefore, the policy shock at the turning point from the second to the third stage of China's family planning initiative may provide a good natural experiment for a valid RD design. Considering the intensive implementation of OCP and a potential 10-month gestation period for all mothers, we predict that there should be a discontinuous jump in the ratio of single-child birth in October 1980. To verify this prediction, we plot in Fig. 1 the relationship between monthly single-child ratios in China and the time lag (in months) between individual birth dates and the assumed cutoff point (October 1980) within the RD bandwidth (approximately 3 years), which shows a noticeable discontinuity in the single-child ratio in our data (the single-child status is directly asked and reported by the Inter-census Survey).

One may argue that OCP can also affect the single-child status of people born within 18–24 months before the policy cutoff point, as their mothers needed the time to recover from the first child-birth and were not physiologically capable of bearing a second child around the policy cutoff. A similar argument can also be made for people born years before the cutoff point and still affected by OCP, as their parents may be psychologically or financially incapable of having a second child in or before October 1980. However, the presence of the above individuals does not change or blur our chosen cutoff point, because these individuals would not be the “compliers” for the RD design (in fact, they should be regarded as the “never takers” under the RD setting). As mentioned, the RD estimation only focuses on the compliant parents who are both willing and able to give birth to another child when OCP was initially carried out, i.e., they should be free to make fertility decisions at the policy cutoff in absence of OCP. For the above reasons, we propose (and will formally prove hereinafter) that there is indeed a plausible cutoff point, October 1980, for our RD design.

⁶ Han, the ethnic majority in China, accounts for more than 90% of the Chinese population.

⁷ To be more specific, the “carrots” for an obedient urban couple include: extra amounts of foodstuff allocation, priority of housing-lot allocation, a monthly cash payment till the single child reaches 15 years old, free nursery and medical care for the child, free primary and secondary schooling for the child, etc. (Banister, 1987; Li and Cooney, 1993). The “sticks”, on the other hand, are the fines imposed on the defiant urban couples: for those who do not hold the one-child certificates, the fines are 3–10 times the yearly per capita disposable income in the correspondent year; for the certificate holders, the fines are even greater, including the repayment of the expected full value of the benefits received under the program (Cooney et al., 1991).

3. Regression discontinuity design based on OCP

The RD method was first introduced by [Thistlethwaite and Campbell \(1960\)](#), and began to be widely applied to economic research in late 1990s. In drawing causal inference between variables, RD serves as a good alternative to the randomized control trials (RCTs), which is traditionally considered the gold-standard solution for endogeneity ([Imbens, 2010](#)). Due to the ethical concerns and high administrative costs, RCTs are seldom used in social policy contexts, and economic researchers often resort to “quasi-experimental” designs for impact evaluations. Among the other common methods such as IV, difference-in-differences and matching techniques, RD designs are deemed to be the “close cousin” of RCTs with the greatest “internal validity” among the alternative quasi-experimental estimators ([Hahn et al., 2001](#); [Lee, 2008](#); [Lee and Lemieux, 2010](#)).

There are two types of RD designs – the “sharp” designs and the “fuzzy” designs. In sharp RD designs, the probability of individuals receiving a treatment raises from 0 to 1 across the cutoff; while in fuzzy RD designs, the probability increases discontinuously but not from 0 to 1. In our context, the sample probability of being a single child jumps from 38% to 43% at the cutoff, which suggests the use of fuzzy RD designs in accordance to many prior studies ([Angrist and Lavy, 1999](#); [Van der Klaauw, 2002](#); [Lei et al., 2010](#); [Chen et al., 2013](#)).

In this paper, we resort to the fuzzy RD framework and estimate the local average treatment effect (LATE) of the OCP on educational attainment indicated by [Eq. \(1\)](#):⁸

$$\tau = \frac{\lim_{X \rightarrow c^+} E[Y|X] - \lim_{X \rightarrow c^-} E[Y|X]}{\lim_{X \rightarrow c^+} E[D|X] - \lim_{X \rightarrow c^-} E[D|X]} \quad (1)$$

For reasons specified in the Technical Appendix, we will mainly rely on the non-parametric approach to estimate [Eq. \(1\)](#), in which we run the one-sided kernel regressions at the cutoff point in X and obtain the point-wise prediction for the four limiting values in [Eq. \(1\)](#) as follows:

$$\lim_{X \rightarrow c^+} E[Z|X] = \frac{\sum_{i: X_i > c} Z_i \cdot K((X_i - x)/h)}{\sum_{i: X_i > c} K((X_i - x)/h)} \quad (2)$$

$$\lim_{X \rightarrow c^-} E[Z|X] = \frac{\sum_{i: X_i < c} Z_i \cdot K((X_i - x)/h)}{\sum_{i: X_i < c} K((X_i - x)/h)} \quad (3)$$

where Z can be replaced with Y or D , $K(\cdot)$ is the kernel function, and h is the bandwidth for the kernel regressions. It is shown that the non-parametric estimate of [Eqs. \(2\) and \(3\)](#) is numerically equivalent to an IV estimate when the uniform kernels is selected for $K(\cdot)$ ([Hahn et al., 2001](#)). Thus the treatment effect in our RD design can also be estimated by a parametric approach through the traditional IV/2SLS method ([Imbens and Lemieux, 2008](#); [Lee and Lemieux, 2010](#)). The estimation equations can be specified as follows:

$$Y = \theta + \tau D + f(X - c) + \varepsilon \quad (4)$$

$$D = \gamma + \delta T + g(X - c) + u \quad (5)$$

As the RD estimates are sensitive to the kernel function choices, we will use the more general non-parametric method for our main results and apply the IV/2SLS estimation (with the uniform kernel specification) for robustness checks. Furthermore, statistical literature shows that the triangle kernel is optimal for the point estimates at boundaries ([Fan and Gijbels, 1996](#); [Lee and Lemieux, 2010](#)). Hence, we use the triangle kernel in our main estimation and the rectangular kernel for robustness checks. For bandwidth choices, we follow [Imbens and Kalyanaraman \(2012\)](#) to use the optimal bandwidth that minimizes the mean squared error (MSE), and present the results based on alternative bandwidths (i.e., one, two, and three years) for robustness checks.

In addition to the abovementioned robustness checks, we also follow the convention of RD literature and perform such validity tests as whether there is extraneous discontinuity in the treatment variable D or the outcome variable Y away from the cutoff point, and whether the density of the assignment variable X exhibits any jumps at the cutoff point. The results of these validity tests and robustness checks will be reported in [Section 5](#).

4. Data

The micro-level data we use are collected from the 2005 Inter-census Survey (also called the 1% Population Sampling Survey) conducted by the National Bureau of Statistics (NBS) of China. The survey is arranged by the State Council of China

⁸ Detailed derivation can be found in the Technical Appendix. In the LATE framework, there are four types of individuals – compliers, never-takers, always-takers and defiers. The basic LATE estimate focuses on the treatment effect among compliers and assumes no defiers in the sample (commonly known as the monotonicity assumption). In our context, it is barely possible that people who decide to raise only one child before OCP would choose to raise more children after OCP, thus the monotonicity assumption is likely to be valid.

Table 2.
Sample summary statistics of main variables.

Variable	Definition	(1) Overall	(2) Pre-OCF Cohort	(3) Post-OCF Cohort
Eduyear	Education (years of schooling)	12.67 (2.872)	12.78 (2.954)	12.54* (2.778)
Yhat	“Compensated” years of schooling	12.69 (2.866)	12.59 (2.954)	12.80* (2.766)
Nosibs	Being a single child (1 = yes, 0 = no)	0.387 (0.487)	0.333 (0.471)	0.444* (0.497)
Healthy	Being healthy (1 = yes, 0 = no)	0.997 (0.057)	0.997 (0.055)	0.997 (0.059)
Married	Being married (1 = yes, 0 = no)	0.384 (0.486)	0.551 (0.497)	0.206* (0.404)
Male	Being male (1 = yes, 0 = no)	0.504 (0.500)	0.521 (0.500)	0.487* (0.500)
N	Sample size	20,174	10,397	9777

Notes: (i) Data source: 2005 China's Inter-census Survey. OCP refers to One-Child Policy. (ii) The table reports the mean values and standard deviations (in parentheses) of the variables for the study sample; (iii) * denotes 10% significance level in the differences between children born before and after the cutoff (October 1980); (iv) the pre-OCF cohort refers to those born within 3 years before the cutoff while the post-OCF cohort refers to those born within 3 years after it; (v) the “compensated” years of schooling is calculated by eliminating from the primitive years of schooling the impacts from the birth year and birth quarter based on the OLS regression coefficients.

and financially supported by the central and local governments, covering about 13 million respondents across the country that are selected by a stratified multistage clustering sampling method from the 2000 Census.⁹ The 2005 Inter-census Survey provides important information on population demography as well as detailed profiles on education, health, employment status, job classification, migrant status, family characteristics, etc.

We obtain a subsample containing 2.6 million observations that are randomly drawn from the whole Inter-Census Survey sample. Since we are interested in the local average effect of OCP, we first narrow the sample down to a smaller subsample in which all observations are born just before and after OCP's introduction in 1979, i.e. they were born during 1975–1985.¹⁰ Second, we exclude ethnic minorities as OCP is not applied to them; rural residents are also excluded because (1) OCP is not strictly implemented in many rural areas (Li, 1995); (2) the large-scale rural land reform (“household responsibility system” reform) occurred at approximately the same time as OCP and would complicate the RD estimation (Almond et al., 2013). Third, we exclude the provinces where the OCP implementation was less strict so that couples were allowed to have two or more children.¹¹ Finally, we exclude the individuals without a traceable migration history (i.e., a history indicating when and where a person migrated across provinces) prior to the survey as we cannot decide on their birthplaces based on the current residential (*hukou*) locations. Our final study sample therefore contains 30,660 individual respondents. In our main RD regressions, the bandwidths are restricted to 1–3 years to better approximate the local average policy effects. Accordingly, we provide in Table 2 the summary statistics for the observations born within the 3-year boundary around the cutoff point (with a sample size of 20,174).

As shown, out of the 20,174 observations, 48.5% of them are in the post-OCF cohort. The human capital level, measured by the (primitive) years of continuous schooling, is 12.54 years for the post-OCF cohort and 12.78 years for the pre-OCF cohort on average.¹² After factoring out the time trend and seasonality in the observed education years, we may calculate each individual's “compensated” years of schooling that controls for the “cohort effect” in educational attainment,¹³ and it is shown that the average compensated years of schooling are 12.80 years for the post-OCF cohort and 12.59 years for the pre-OCF cohort. More importantly, as Fig. 2 demonstrates, the average years of schooling jump upwards at the cutoff point, indicating a discontinuity in the distribution of human capital for people born right before and after the OCP.

⁹ In China, the full population census is conducted every 10 years, and the inter-census 1% population sample survey is conducted between two consecutive census surveys. The inter-census survey follows a three-stage clustering sampling method, in which the first stage is set at the provincial level and the ultimate sampling unit is set at the residential community level (enumeration districts).

¹⁰ Since we assume that the cutoff point is October 1980, we narrow the observations down to those who were born between October 1975 and October 1985. In the RD regressions, we allow bandwidths to change from 1 to 3 years, which allows us to move the cutoff point up to 2 years away from the assumed one in the validity tests.

¹¹ Based on Guo (2003), the less restricted provinces are Henan, Guangxi, Gansu, Hebei, Inner Mongolia, Yunnan, Qinghai, Ningxi, Hainan, Xinjiang, etc.

¹² Since the individuals in our sample are relatively young (around 25 years old), those in the younger (post-OCF) cohort are more likely to go back to school for further education. For example, they may temporarily leave school to accumulate work experience with the plan to pursue a higher degree (such as an MBA) a few years later. In this sense, we admit that the 2010 Census or 2015 Inter-census data might be better for our analysis, but we do not have access to these data at the time this study was carried out.

¹³ To control for the time trend and seasonality of education as suggested by Buckles and Hungerman (2013) and Angrist and Krueger (1991), we eliminated from the observed years of schooling the impacts of the birth year and quarter dummies based on the OLS coefficient estimates. That is, after estimating $eduyear = \beta_0 + \beta_1 \cdot year + \beta_2 \cdot quarter + u$ by an OLS regression, we calculate the “compensated” years of schooling by the formula $yhat = eduyear - b_1 \cdot year + b_2 \cdot quarter$ (where b is the OLS estimate of β). Furthermore, since we are interested in people's permanent education attainment, we evaluate the years of schooling for those still in school by the maximum years needed to achieve their ongoing education level.

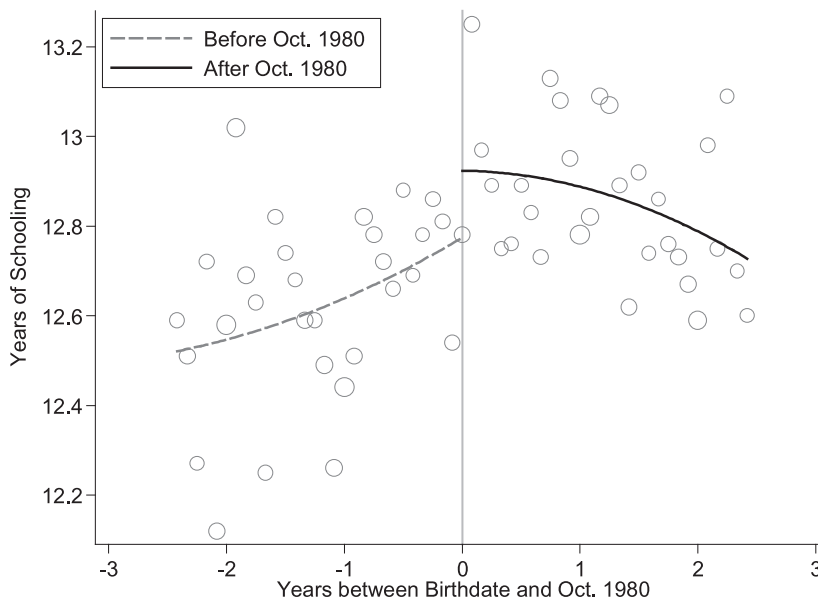


Fig. 2. Discontinuity in the years of schooling (outcome variable) around Oct. 1980. Notes: (i) Data source: 2005 China's Inter-census Survey. (ii) The hollow circle represents the “compensated” average years of schooling (after detrending and deseasonalizing the observed years of schooling) for the children born in the corresponding month, and its size corresponds to the sample size; (iii) the solid line represents the quadratic predictions of the relationship between the outcome variable (years of schooling) and the assignment variable (adjusted month) after October 1980, and the dashed line represents the relationship before October 1980.

Table 2 indicates that the average single-child ratio for the post-OCF cohort is significantly higher than that for the pre-OCF cohort (44.4% vs. 33.3%), which in turn signifies the tightening of the family planning campaign during the sample period. For the validity tests and robustness checks, we use a set of baseline covariates in the RD estimation, which includes gender, health, marital status and province of birth.¹⁴ As Table 2 shows, the average rate of being married for the younger cohort is much lower than that for the older one (20.6% vs. 55.1%), which indicates that the probability of getting married increases with age; the average sex ratio (defined as the ratio of male to female in the sample) is higher for the older group (1.09 vs. 0.95), which is in accordance to the 2005 China inter-census's aggregate-level statistics published by NBS of China¹⁵ and thus indicates a lower male representation in the post-OCF cohort¹⁶; although the “education-health gradient” is verified in the previous literature (Behrman, 1996; Glewwe et al., 2001; Smith, 2004), there is no strong evidence that the average health status is different between the pre- and post-OCF cohorts in our sample, which in turn suggests that the jump in the average years of schooling at the cutoff point may not be a result of the change in health status (this finding will be tested formally by regressions in Section 5).

5. Estimation results

5.1. The causal effect of OCF on educational attainment

In the previous sections, we graphically showed the discontinuities of the single-child ratio and the years of schooling at the policy cutoff point (October 1980). To verify the jumps at the cutoff, we run the kernel regressions using the triangle kernel based on Eqs. (2) and (3), and report the results in Table 3. As shown, the estimated jumps in the outcome and the treatment are the non-parametric estimates for the numerator and the denominator in Eq. (1), respectively. By applying the procedure suggested by Nichols (2007), we estimate the LATE of OCF on the individuals' educational attainment. The results suggest that OCF instantly increases the tendency of being a single child among the respondents born around the cutoff point with a significant 4.8% jump in probability, and the decline in this probability significantly increases the respondents' educational attainment by 0.27 years, indicating that the LATE estimate is 5.6 years of schooling.

The results are robust when we change the bandwidth from 1 to 3 years, as indicated by the different columns of Table 3. Therefore, the above results strongly suggest that being a single child significantly increases one's long-term human capital

¹⁴ Due to the relatively young age for the sample individuals, marriage can interrupt their “continuous” schooling. The provincial dummy variables are constructed according to the GB/T2260 code of China, and the health dummy variable is constructed in the same way as Lei et al. (2010).

¹⁵ The statistics can be found in Table 3-1a on <http://www.stats.gov.cn/tjsj/ndsj/renkou/2005/renkou.htm>.

¹⁶ Although Hesketh et al. (2005) conclude that the sex ratio has been steadily increasing since the onset of OCF, the average sex ratio is indeed lower among the urban residents aged 21–24 (the post-OCF cohort) than the urban residents aged 25–28 (the pre-OCF cohort) in 2005 based on the aggregate data from the NBS of China.

Table 3.

Non-parametric estimates for the effects of OCP using different bandwidths.

Estimator	(1) 2.4 Years	(2) 1 Year	(3) 2 Years	(4) 3 Years
Jump in eduear	0.271*** (0.0992)	0.393*** (0.153)	0.252** (0.110)	0.281*** (0.0891)
Jump in nosibs	0.048*** (0.0172)	0.069*** (0.0268)	0.042** (0.0191)	0.056*** (0.0154)
Weighted LATE	5.648** (2.405)	5.676** (2.591)	6.026* (3.151)	5.020*** (1.751)
Cutoff point	Oct. 1980	Oct. 1980	Oct. 1980	Oct. 1980
Sample size	16,269	7037	14,814	20,174

Notes: (i) Data source: 2005 China's Inter-census Survey. OCP refers to One-Child Policy. (ii) The “jump” is the estimated difference between the pre- and post-OCP cohorts at the cutoff point, i.e., the numerator or denominator in Eq. (1) in the main text; (iii) bootstrap standard errors are in parentheses; (iv) ***, ** and * denote 1%, 5%, and 10% significance levels, respectively; (v) the non-parametric estimates are obtained from kernel regressions using triangle kernel without incorporating the baseline covariates; the first column corresponds to the optimal bandwidth (2.4 years) following Imbens and Kalyanaraman (2012); (vi) we use 30,660 observations (born 5 years before and after the cutoff point) to draw the optimal bandwidth; (vii) the definitions of “eduear” and “nosibs” are provided in Table 2.

accumulation, which is in accordance to the classic Q-Q trade-off theory. According to Becker (1960) and other related works, each couple independently chooses the optimal number of children and optimal human capital investment in each child by maximizing their utility functions when the restrictions on fertility and investment do not bind. For the “compliers” of OCP, their optimal fertility choices are effectively bound by the policy from two perspectives. First, the policy increases the shadow price of having a second child by imposing heavy fines, making it less cost-effective for them to have more than one child. Second, the policy also decreases the income elasticity of child quantity since couples who hold the one-child certificate can keep gaining benefits in the long run, making it less appealing for them to violate OCP. As a result, a binding constraint on the number of children would prevail in the compliers' utility maximization problem, and their optimal choices will become corner solutions as they are compelled to transfer all or part of the parent-to-offspring investments from later-born children into the first child.¹⁷

Several cautions should be taken when interpreting the above results. First, it should be noted that the above policy effect is a local average estimate that reflect the OCP impact among the compliant families, and the quantitative estimate cannot be directly generalized to the whole population. One important reason is that the compliers are likely to have different characteristics compared to others, because the punishment for violating the one-child policy might vary across socioeconomic groups. For instance, self-employed individuals were usually punished in monetary term only, while government officials could be denied the opportunity of promotion or even lose their positions, making them less likely to violate the policy. As mentioned, RD has the greatest “internal validity” among the quasi-experimental estimators while its “external validity” is limited to certain extent, so we must be very careful when trying to extrapolate the LATE estimate results to others. Second, it should also be noted that the above policy effect is estimated by comparing the single children with the non-single ones, thus it reflects a “cumulative” treatment effect of OCP, i.e., the overall impact on educational attainment by reducing the quantity of children per couple from two, three or even more to one single child; apparently, this cumulative effect is conceptually different from the “marginal” effect of one-child reduction in the family. Third, due to the quasi-experimental nature of the RD design, we are reasonably confident that the above estimate of treatment effect represents a causal relationship between the quantity and quality of children among the sample respondents, as the endogeneity of family fertility choices is removed by the external OCP shock (indicated by the discontinuous jump in the single-child ratio). The validity tests and robustness checks (including a falsification test) for the causal inference will be discussed below.

5.2. Validity tests for the RD design

In the RD framework, several important assumptions need to be made in order to draw the causal inference between the treatment variable and the outcome variable (Nichols, 2007), and thus it is imperative to verify the validity of our results by testing whether each assumption holds.

The first assumption of RD designs is that the outcome and treatment variables both jump discontinuously at the cutoff point of the assignment variable. This is the most basic assumption, without which we cannot identify the causal relationship between the two variables. In fact, the validity of this assumption was already verified in the above kernel regressions. As already shown in Table 3, the estimated jumps are all significant on the 5% (or even 1%) confidence level.

Second, we need to make sure that the outcome and treatment variables are both continuous functions of the assignment variable in absence of the policy treatment. To verify this assumption, we can test if there are other extraneous

¹⁷ If we set the parental investment (including spending during pregnancy) in the first child as one good and the investment in the later-born children as the other good, then the latter input will become 0 (corner solution) for the compliers when OCP is in place. Apparently, the optimal choices will be interior solutions in absence of OCP.

Table 4.

Test for discontinuity of outcome and treatment at alternative cutoff points.

Estimator	(1) Oct. 1979	(2) Apr. 1980	(3) Apr. 1981	(4) Oct. 1981
Jump in edueyear	0.099 (0.090)	0.117 (0.098)	−0.041 (0.135)	−0.014 (0.113)
Jump in nosibs	0.027* (0.015)	0.012 (0.017)	0.035 (0.023)	0.009 (0.020)
Weighted LATE	3.615 (3.332)	9.810 (13.55)	−1.176 (4.160)	−1.539 (13.63)
Sample size	19,536	16,890	9158	10,478

Notes: (i) Data source: 2005 China's Inter-census Survey. (ii) The “jump” is the estimated difference between the pre- and post-OCP cohorts at the cutoff point, i.e., the numerator or denominator in Eq. (1) in the main text; (iii) bootstrap standard errors are in parentheses; (iv) ***, ** and * denote 1%, 5%, and 10% significance levels, respectively; (v) the kernel and optimal bandwidth choices are the same as Table 3; (vi) all regressions are without the baseline covariates; (vii) we use 30,660 observations (born 5 years before and after the cutoff point) to draw the optimal bandwidth; (viii) the definitions of “edueyear” and “nosibs” are provided in Table 2.

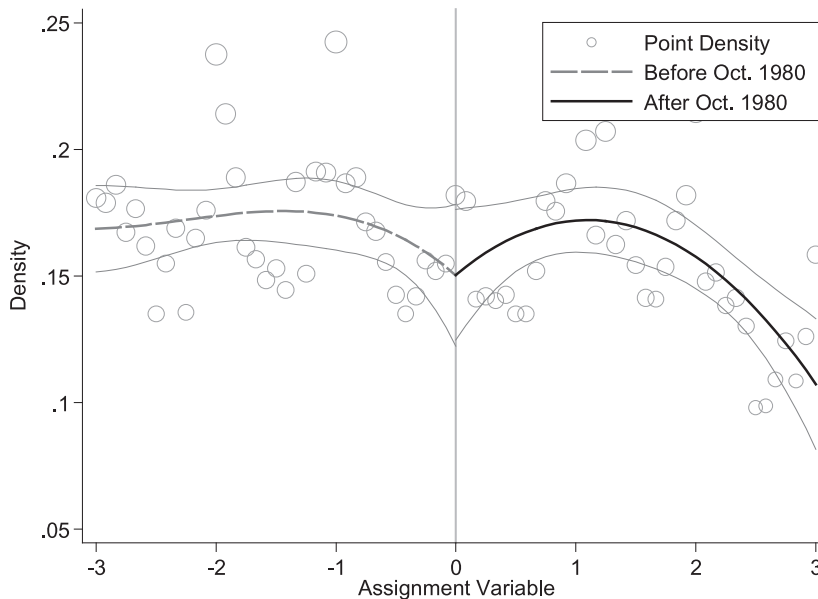


Fig. 3. McCrary test for the discontinuity of assignment variable. Notes: (i) Data source: 2005 China's Inter-census Survey. (ii) The hollow circle represents the estimated point density at the corresponding value of the assignment variable, and its size corresponds to the sample size; (iii) the solid line represents the fractional polynomial estimation for the density of the assignment variable after October 1980, and the dashed line represents the density before October 1980; (iv) the estimations on both sides are associated with the 95% confidence bands.

discontinuities in the outcome and treatment variables away from the presumed cutoff point (October 1980). We therefore run the same RD regressions for some other hypothetical cutoff points within a one-year boundary of October 1980. As shown in Table 4, the outcome and treatment variables are generally continuous at these alternative points, which diagnostically verifies that Assumption 2 does hold.

Moreover, we need to ensure that individuals cannot precisely manipulate the assignment, so they cannot be sorted across the cutoff in response to the policy treatment, and also the baseline covariates should be continuous at the policy cutoff. If individuals could precisely manipulate the assignment,¹⁸ then the compliers would have been born earlier and therefore we should see more observations clustering in the time period just before the cutoff point and that the density of the assignment variable would be discontinuous.¹⁹ Following the literature convention, we apply the McCrary (2008) test to verify this no-manipulation assumption. The estimated log difference between the density just before and just after the cutoff point is about 0.012 (with its standard error being 0.042), which suggests that individuals cannot precisely control the assignment by crowding in. The graphical presentation for the McCrary test below (Fig. 3) also verifies this assumption.

¹⁸ Since children are not supposed to make decisions before being born, the manipulation refers to the decision of their parents in giving birth before the implementation of OCP in order to avoid punishment.

¹⁹ This would be the case if some parents who initially plan to have their second child after October 1980 decided to give birth to the child before the policy cutoff (e.g., through Caesarean section).

Table 5.

Test for discontinuity of baseline variables at the cutoff point.

Estimator	(1) Opt. Bandwidth	(2) 1 Year	(3) 2 Years	(4) 3 Years
Jump in male	–0.031 (0.023)	–0.028 (0.027)	–0.019 (0.020)	–0.000 (0.016)
Jump in married	0.039 (0.023)	0.043 (0.026)	0.022 (0.019)	0.001 (0.015)
Jump in healthy	0.002 (0.003)	–0.000 (0.003)	–0.001 (0.002)	–0.001 (0.002)
Jump in provincial category dummy variables	All insignificant			

Notes: (i) Data source: 2005 China's Inter-census Survey. (ii) The table reports the estimated differences of the characteristics between the pre- and post-OCP cohorts at the cutoff point; (iii) bootstrap standard errors are in parentheses; (iv) ***, ** and * denote 1%, 5%, and 10% significance levels, respectively; (v) the cutoff point (October 1980), kernel and optimal bandwidth choices are the same as Table 3; (vi) the definitions of “male”, “married”, and “healthy” are provided in Table 2, while the provincial category dummy variables are a group of binary variables indicating the birth provinces; (vii) each cell refers to the result of a different regression.

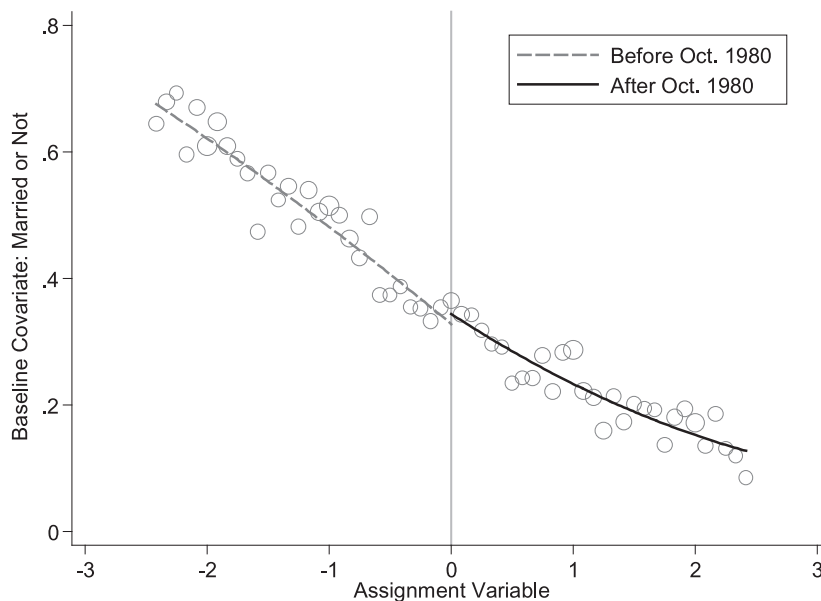


Fig. 4. Continuity of the baseline covariate (marital status). Notes: (i) Data source: 2005 China's Inter-census Survey. (ii) The hollow circle represents the average percentage of being married in each month, and its size corresponds to the sample size; (iii) the solid line represents the quadratic predictions of the relationship between the baseline covariate (dummy indicator of being married) and the assignment variable after October 1980, and the dashed line represents the relationship before October 1980.

Furthermore, under the assumption, we should also see that the predetermined characteristics exhibit no jumps at the cutoff point. According to Lee and Lemieux (2010), if the compliers could precisely manipulate the assignment in order to get a different outcome, they would change their *ex-ante* characteristics related to the outcome and thus those baseline covariates would jump at the cutoff point. To verify this prediction, we run the RD kernel regressions with the outcome variable being replaced by the baseline covariates. As shown in Table 5, all the covariates are continuous at the cutoff point, which suggests that the policy effect on the educational attainment is not driven by the change in the individuals' baseline characteristics other than their single-child status.²⁰

To illustrate the above test graphically, we take the respondents' marital status as an example and draw the sample distribution of this covariate in Fig. 4. The graph clearly shows that the probability of getting married decreases continuously along with the reduction of age (assignment variable), and there is no evident jump at the cutoff point.

²⁰ Admittedly, there may be other variables that can affect individuals' education but are not included in Table 5. Although it is impossible to exhaust all the possibilities, we nevertheless also check the continuity of the following variables using similar RD regressions and find no evident jumps at the cutoff point: the enrollment status for the medical insurance and unemployment insurance, the number of pregnant times (only for female respondents with pregnancy histories), the household type, the contract worker status, and the weekly working hours.

Table 6.

Non-parametric estimates for the effects of OCP (with baseline covariates).

Estimator	(1) 2.2 Years	(2) 1 Year	(3) 2 Years	(4) 3 Years
Jump in edueyear	0.270** (0.123)	0.271* (0.153)	0.271** (0.125)	0.238** (0.117)
Jump in nosibs	0.060*** (0.021)	0.067*** (0.026)	0.060*** (0.021)	0.055*** (0.020)
Weighted LATE	4.505** (2.232)	4.023* (2.372)	4.506** (2.257)	4.311* (2.279)
Baseline covariates	Yes	Yes	Yes	Yes
Cutoff point	Oct. 1980	Oct. 1980	Oct. 1980	Oct. 1980
Sample size	15,274	7037	14,814	20,174

Notes: (i) Data source: 2005 China's Inter-census Survey. OCP refers to One-Child Policy. (ii) The "jump" is the estimated difference between the pre- and post-OCP cohorts at the cutoff point, i.e., the numerator or denominator in Eq. (1) in the main text; (iii) bootstrap standard errors are in parentheses; (iv) ***, ** and * denote 1%, 5%, and 10% significance levels, respectively; (v) the covariates include all baseline variables in Table 5, and also the yearly and seasonal dummies; (vi) the kernel and optimal bandwidth choices are the same as Table 3; (vii) we use 30,660 observations (born 5 years before and after the cutoff point) to draw the optimal bandwidth; (viii) the definitions of "edueyear" and "nosibs" are provided in Table 2.

Table 7.

Non-parametric estimates for the effects of OCP (using rectangular kernel).

Estimator	(1) 1.9 Years	(2) 1 Year	(3) 2 Years	(4) 3 Years
Jump in edueyear	0.301*** (0.104)	0.251* (0.139)	0.361*** (0.099)	0.330*** (0.081)
Jump in nosibs	0.053*** (0.018)	0.052** (0.024)	0.067*** (0.017)	0.077*** (0.014)
Weighted LATE	5.634** (2.249)	4.805* (2.897)	5.373*** (1.669)	4.277*** (1.096)
Cutoff point	Oct. 1980	Oct. 1980	Oct. 1980	Oct. 1980
Sample size	12,739	7037	14,814	20,174

Notes: (i) Data source: 2005 China's Inter-census Survey. OCP refers to One-Child Policy. (ii) The "jump" is the estimated difference between the pre- and post-OCP cohorts at the cutoff point, i.e., the numerator or denominator in Eq. (1) in the main text; (iii) bootstrap standard errors are in parentheses; (iv) ***, ** and * denote 1%, 5%, and 10% significance levels, respectively; (v) the non-parametric estimates are gained from kernel regressions using rectangular kernel without incorporating the baseline covariates; (vi) the rule of choosing the optimal bandwidth is the same as Table 3; (vii) we use 30,660 observations (born 5 years before and after the cutoff point) to draw the optimal bandwidth; (viii) the definitions of "edueyear" and "nosibs" are provided in Table 2.

5.3. Additional robustness checks

This sub-section contains further robustness checks in addition to the traditional validity tests for the RD estimation. First of all, although it is not necessary to control the baseline covariates to obtain consistent estimates of the treatment effect if the no-manipulation assumption holds, we nevertheless do it to check the robustness of our results. As shown in Table 6, the estimates are all statistically significant and almost identical to those in Table 3.

Second, as discussed above, the triangle kernel is optimal for non-parametric estimates at the boundary. However, it would be more reassuring if the results are robust to the variation of kernel functions. Table 7 reports the results using the rectangular kernel in the non-parametric regressions. As indicated, the results are generally significant and similar to those in Table 3.

Third, we show that the LATE estimates can also be obtained through a parametric approach. As suggested by Hahn et al. (2001); Lee and Lemieux (2010) and several other studies, we may implement the IV/2SLS estimation under the RD framework following the steps in Eqs. (4) and (5). Table 8 reports these estimates. As expected, the IV/2SLS estimates of the effect of being a single-child on educational attainment are all statistically significant and quantitatively consistent with the non-parametric results. This in turn suggests that our main non-parametric results are stable across alternative and equally plausible functional specifications.

Table 8 also reports the OLS estimates of the treatment effect. Comparing the OLS results with the IV/2SLS results, we find that ignoring the endogeneity of individuals' single-child status can lead to an underestimation of the OCP treatment effect. There are at least two reasons: first, the underestimation may arise from the reverse causality, i.e. a better educational outcome of the first child can encourage parents to have more children; second, some unobserved factors can be correlated to both the odds of being a single child and the educational attainment, e.g., parents with (unobserved) strong son-preference and who happen to have their first child as a son may choose not to have more children and also to invest more resources in their son's education. The potential problems (e.g. estimation bias) that associate with the above endogeneity thus justify the use of quasi-experimental estimation such as the RD design.

Table 8.

OLS and IV/2SLS estimates for the treatment effects of OCP.

	(1) 2.4 Years	(2) 1 Year	(3) 2 Years	(4) 3 Years
OLS				
Without covariates	1.880*** (0.042)	1.823*** (0.064)	1.855*** (0.046)	1.843*** (0.038)
With covariates	1.649*** (0.046)	1.597*** (0.068)	1.636*** (0.049)	1.619*** (0.042)
IV/2SLS				
Without covariates	4.607*** (1.755)	5.237* (3.031)	5.511*** (1.638)	4.257*** (0.999)
With covariates	5.439* (3.147)	4.092* (2.422)	6.337* (3.416)	5.646* (3.254)
Sample size	16,269	7037	14,184	20,174

Notes: (i) Data source: 2005 China's Inter-census Survey. OCP refers to One-Child Policy. (ii) The table reports the LATEs estimated by the parametric approach (OLS vs. IV/2SLS); (iii) robust standard errors are in parentheses; (iv) ***, ** and * denote 1%, 5%, and 10% significance levels, respectively; (v) the reported parameter estimates are for the treatment effect of being a single child on years of schooling in an OLS or 2SLS estimation; (vi) each column is associated with a bandwidth for the corresponding sample selection; (vii) the baseline covariates are the same as in Table 6; (viii) each cell refers to the result of a different regression.

Table 9.

Falsification test on the effects of OCP (using non-Han sample).

Estimator	(1) 2.3 Years	(2) 1 Year	(3) 2 Years	(4) 3 Years
Jump in eduyear	0.014 (0.208)	0.030 (0.324)	0.009 (0.227)	−0.044 (0.183)
Jump in nosibs	−0.008 (0.023)	−0.023 (0.035)	−0.013 (0.0250)	−0.007 (0.020)
Weighted LATE	−1.590 (25.98)	−1.319 (14.63)	−0.678 (18.31)	6.095 (26.37)
Cutoff point	Oct. 1980	Oct. 1980	Oct. 1980	Oct. 1980
Sample size	4743	2306	4609	6717

Notes: (i) Data source: 2005 China's Inter-census Survey. OCP refers to One-Child Policy. (ii) The “jump” is the estimated difference between the pre- and post-OCP cohorts (non-Han) at the cutoff point, i.e., the numerator or denominator in Eq. (1) in the main text; (iii) bootstrap standard errors are in parentheses; (iv) ***, ** and * denote 1%, 5%, and 10% significance levels, respectively; (v) the kernel and optimal bandwidth choices are the same as Table 3; (vi) all regressions are without the baseline covariates; (vii) we use 10,035 observations (born 5 years before and after the cutoff point) to draw the optimal bandwidth; (viii) the definitions of “eduyear” and “nosibs” are provided in Table 2.

Last but not least, we implement a falsification (placebo) test based on the excluded sample to validate our findings on the study cohorts. One potential complication factor for our study is China's higher education reform in 1999 (also dubbed the “great leap forward in higher education”) that results in a significant expansion in college enrollment among people born after 1980 (Li and Xing, 2010). To verify that the observed jump in educational attainment of the study cohorts is not caused by such non-OCP events, we apply the RD estimation on the excluded sample of non-Han people living in the same provinces as the main sample respondents. The reason for selecting the ethnic minorities is that they are not restricted by OCP but are directly affected by the 1999 education reform in China.²¹ Table 9 reports the RD regression results for this placebo test, which shows that neither the outcome variable nor the treatment variable experiences a discontinuous jump at the cutoff point. Since the ethnic minorities were given preferential treatment for college entrance in the 1999 higher education reform,²² the potential impact of the reform on educational attainment should be arguably larger among this group compared to the Han people. Thus, the above non-significant results imply that the previously observed increase in the years of schooling among the Han cohorts is not likely to be caused by the 1999 college expansion reform, which in turn verifies our hypothesis on the OCP effect.

In summary, the above validity/robustness tests confirm that OCP has a significant and positive effect on the human capital stock among the single children in the compliant families, and that there exists a negative causal relationship between the quantity and quality of children.

²¹ Since OCP was never imposed upon the non-Han couples, the ethnic minorities in the corresponding provinces form an ideal sample for the placebo test as they are not subject to the OCP influence in these regions (note that we also exclude the unemployed, migrants, and school dropouts from the ethnic minority sample).

²² According to the landmark policy document “Action Scheme for Invigorating Education towards the 21st Century” published by the State Council of China in 1999, the educational attainment of the ethnic minorities were given priority in the education reform agenda, although this group already enjoyed preferential policies for college entrance (compared to the Han people) before 1999.

Table 10.
Subsample non-parametric estimation results for the effects of OCP.

Estimator	(1) Same-Sex	(2) Sons	(3) Daughters	(4) Rich	(5) Poor
Jump in eduyear	0.393*** (0.114)	0.561*** (0.172)	0.204 (0.156)	0.134 (0.142)	0.349*** (0.128)
Jump in nosibs	0.071*** (0.020)	0.059** (0.029)	0.082*** (0.030)	0.046* (0.027)	0.049** (0.020)
Weighted LATE	5.500*** (1.910)	9.548** (4.863)	2.504 (1.921)	2.886 (2.999)	7.132** (3.375)
Sample size	16,535	8750	7785	15,591	15,069

Notes: (i) Data source: 2005 China's Inter-census Survey. OCP refers to One-Child Policy. (ii) The "jump" is the estimated difference between the pre- and post-OCP cohorts within each subgroup at the cutoff point, i.e., the numerator or denominator in Eq. (1) in the main text; (iii) bootstrap standard errors are in parentheses; (iv) ***, ** and * denote 1%, 5%, and 10% significance levels, respectively; (v) the non-parametric estimates are obtained from kernel regressions using the triangle kernel without incorporating the baseline covariates; (vi) the rule of choosing the optimal bandwidth and the cutoff point (October 1980) are the same as Table 3; (viii) the definitions of "eduyear" and "nosibs" are provided in Table 2.

6. Subsample analyses

As discussed in the Introduction, the Q-Q trade-off theory is established upon three key presumptions: impartial parents, binding budget constraints and credit market failures. If the presumptions do not hold, the observed negative relationship between the quantity and quality of children can disappear. Although it might be impossible to confirm some of the presumptions, we can still verify them in a somewhat diagnostic way by the following subsample analyses.

6.1. Are parents partial to boys?

In the classical Q-Q trade-off theory, we presume that parents are impartial to any of their children. However, it is widely believed that sons are preferred to daughters in many families in China. This violation of the presumption can hide the negative relationship from being observed: assume that we have one "single" child and one "non-single" male child drawn from a big family, say, who has three sisters; if the parents are indeed partial to boys, they may put most of the human capital investments on the boy rather than his sisters, and therefore, when we compare these two observations in our regressions, we may not see an evident difference between their levels of human capital. To see whether this is an adverse factor behind our estimation, we run the RD regressions (non-parametrically) on a subsample of families with the "same-sex children" by excluding the respondents with sibling(s) of different genders.

As shown in the first column of Table 10, the significance levels for this subsample estimation (the LATE estimate) are upgraded from 5% to 1% compared to those in the first column of Table 3. Moreover, the discontinuous jumps in the outcome and the treatment among the "same-sex children" families are much greater than those among the "mixed" ones (e.g., 0.393 vs. 0.271 for the jump in education). These results strongly support the abovementioned hypothesis: if parents are indeed partial to boys, then the significance levels of the RD estimates should increase when we exclude the complication factor from regressions (i.e., the "same-sex children" families should not have the gender-related imbalance in the intra-household investment and thus should have more pronounced Q-Q trade-off).

There can be another complication in the presence of gender discrimination: if parents are more willing to invest in the education of boys than girls (i.e., parents value the human capital of boys more than that of girls), the perceived marginal return of educational investment for girls will be lower than that for boys even though the real return is not²³; thus, the educational investment in girls might be lower than the true optimal value, which can result in a vague relationship between quantity and quality among daughters. To see whether this is the case, we further classify the sample by gender. In the male subsample, the respondents are the sons in their respective families, while the female subsample represents the daughter group. As shown in Table 10, the results for the sons group are all significant (with large LATE estimates), while for the daughters the jump in their educational attainment is not evident and thus the estimated LATE is not significant.²⁴ The results again support the popular belief of "son preference" among the Chinese families.

6.2. Is less budget constraint associated with less Q-Q trade-off?

We also presume that parents only have limited resources to invest in their children's human capital when they cannot easily get a loan from the market. However, with the rising household income and the development of credit markets,

²³ The perceived marginal return may not be the actual marginal return, and it only reflects what parents choose to believe. If parents are partial to boys, they may undervalue the marginal return of educational investment in girls.

²⁴ Admittedly, the insignificant estimates can be due to the fact that some parents were allowed to have more than one child if their first child was a girl after the onset of OCP. We cannot completely eliminate this possibility because of the limited hukou information in our data; however, it will not nullify the validity of the causal inference since in fuzzy RD designs individuals in the treatment group are not necessarily affected by the policy.

the binding budget constraints can be relaxed and the Q-Q trade-off can be obscured: an increase in the quantity of children may not translate to a reduction in the parent-to-offspring investments due to their improved borrowing capability. As a result, we may find no evident trade-off (or even a positive correlation) between the quantity and quality of children. To check on this, we carry out another subsample analysis based on income stratification: the “rich” group contains individuals in provinces with per-capita GDP greater than 5000 *Yuan*, and the “poor” group is from the provinces with less per-capita GDP.²⁵ We assume that the parents in richer provinces are subject to less budget constraint due to the overall better economic conditions and the relatively well developed credit markets in these areas as compared to the other sample provinces.

The last two columns of Table 10 present these subsample results. As shown, the difference in educational attainment between the pre- and post-OCP cohorts is not evident in the “rich” group, neither is the estimated LATE, which suggests that the negative relationship between the quantity and quality of children is somehow obscured among the rich; on the contrary, the results for the poor group are all significant even when the sample size becomes smaller, which suggests that the Q-Q trade-off is quite evident among the economically less developed areas. These results to some extent speak to the abovementioned hypothesis: in richer provinces, the relaxed constraint on external borrowing and intra-household human capital investment weakens one of the important presumptions of the Q-Q trade-off theory and thus makes the negative relationship vague.

7. Conclusions

This paper re-examines the classic Q-Q trade-off theory using an innovative RD framework. Economists have been trying to unravel the relationship between the quantity and quality of children for more than half a century since the emergence of “new home economics” in the early 1960s. Nevertheless, the empirical studies aiming to make causal inferences are still far from sufficient. To the best of our knowledge, this paper is the first attempt in this direction through the application of an RD framework. Based on the special historical context of contemporary China and a nationally representative inter-census survey data, we find a good natural experiment for the RD design and document the negative causal relationship between the quantity and quality of population. Our findings not only speak to the classic Q-Q trade-off theory, but also facilitate a better understanding of the intra-household investment in human capital, which in turn may be implicational for the future population policies in China as well as other countries.

With the disproportionate growth of aging population in China and the rising outcry against the country’s birth control policy, the relaxation of OCP has become one of the most discussed issues in recent years, and the “single parent, double children”²⁶ policy implemented after 2013 signals a major change of the national family planning strategy and thus draws massive attention from the public. Some strongly support the relaxation by claiming that the strict family planning will exacerbate the aging burden and have a negative impact on the economic development through the increasing old-age dependency ratio. Others claim that the population control policy is not to be blamed, as the population aging can be a natural result of the “demographic transition” that has accompanied the economic development in many other industrialized countries. No matter which side is correct, one thing is for sure: China is getting old before getting rich. Although China has become the World’s second largest economy in terms of the aggregate GDP, its per-capita GDP is still considerably lower than most OECD countries. Moreover, the systems of pension, social security, health care and education are still rather immature in China. In the current context, if the family planning is drastically relaxed nationwide, a new baby boom may be expected, and whether it will hinder or promote the long-term economic growth depends on whether the increase in young population will be accompanied with a decelerating or accelerating accumulation of per capita human capital. Our RD estimates suggest that, under the special historical background of OCP, a reduction in family size is likely to promote the children’s long-term human capital stock (measured by the years of schooling). One plausible explanation is that the income elasticity of child quantity was at a relatively high level in China during the early stage of economic development, and the implementation of OCP may have forced this elasticity to decrease by extraneously raising the implicit and explicit costs of having more children. As a result, the income elasticity of child quality also increased, and parents may have higher incentives to invest in their children’s education with the improvement of economic conditions. In that sense, the much blamed family planning policy in China may have helped the country improve the quality of its population, which in turn is beneficial to its sustainable economic growth in the long run. An important extension of this finding is that the current relaxation of population control may have the potential to lower (or at least to slow down the accumulation of) the per capita human capital among the younger-aged cohorts, and therefore special cautions should be taken when the relaxation policy is to be implemented across the country.

The subsample analyses in our study also shed light on the potential mechanisms of the Q-Q trade-off by diagnosing its theoretical presumptions. As suggested by the results, the negative Q-Q relationship is more pronounced in the economically less developed regions where the household budget is more constrained and the credit market is less mature. Hence, in order to alleviate the potential negative impact of a population de-control, China should make an effort to enhance the

²⁵ We use the average provincial per-capita GDP in 1993–1995 for the group classification. Since this is the period when China implemented the 9-year compulsory education law, we believe the provincial GDP at this time is a good proxy for the family income related to educational investment.

²⁶ The so-called “single parent, double children” policy allows a couple to have two children if one of them is a “single child” (i.e., with no siblings). In late 2015, the policy was further relaxed to a two-child policy that allows all couples to have two children.

development of the credit markets in the underdeveloped regions to assist the parents to invest in their children's human capital. For example, commercial student loans and the government sponsored tuition subsidy programs can be established and promoted in these areas to help the poor families. In addition, the relaxation of population control should be more prudent in the underdeveloped areas, with the enforcement pace and intensity adjusted according to the local conditions. Lastly, the subsample analyses also reveal the strong son-preference by the Chinese parents, which can also hamper the intra-household investments. Therefore, policies that promote gender equality are also needed to expand the education opportunities for girls, which in turn may facilitate the overall improvement of human capital accumulation in China.

Acknowledgment

This study is supported by National Natural Science Foundation of China (Grant No. 71573003), Ministry of Education of China (Grant No. 12JZD036), Beijing Social Science Foundation Research Base Grant (Grant No. 16JDLJB001), and Peking University School of Economics Research Seed Grant. We thank the National Bureau of Statistics of China for providing the data. We are also grateful to the anonymous referees for their constructive suggestions. Nevertheless, the authors are responsible for all remaining errors.

Technical Appendix for the empirical strategy

In a traditional fuzzy RD design, there are three key variables – assignment variable (also called “forcing variable” or “running variable”), treatment variable, and outcome variable. First, the assignment variable X (defined as the time lag in months – the smallest unit to measure birth time in the dataset – between the individual's date of birth and the cutoff point, October 1980) “assigns” an individual into either the treatment group ($T = 1$, indicating the individual is born after Oct. 1980) or the reference group ($T = 0$, indicating the individual is born before Oct. 1980). Second, the treatment variable D is the dummy variable indicating whether an individual is affected by the treatment policy, in our context whether being a single child in the family ($D = 1$ for yes and $D = 0$ for no). As the nature of fuzzy RD, individuals in the treatment group are not necessarily determined by the treatment policy (i.e., $D \neq T$), while the probability of being affected should be higher in the treatment group than in the reference group. Third, the outcome variable Y is the individual's human capital stock (measured by the years of schooling) that can be affected by discontinuous changes of the treatment variable. Under Rubin's potential outcome framework, Y can be categorized into two types in our context – Y_1 is the outcome for single children and Y_0 is the outcome for non-single children. Hence, Y can also be expressed as:

$$Y = Y_1 \cdot D + Y_0 \cdot (1 - D) = Y_0 + (Y_1 - Y_0) \cdot D \quad (\text{A1})$$

Under the “homogeneity assumption” that individuals in each group have the same level of expected educational attainment, the treatment effect τ of OCP is simply the difference of the expected outcomes between the single and non-single children as shown below:

$$\tau = Y_{11} - Y_{j0} = E[Y_1|X] - E[Y_0|X] = \lim_{X \rightarrow c^+} E[Y|X] - \lim_{X \rightarrow c^-} E[Y|X] \quad (\text{A2})$$

where $\lim_{X \rightarrow c^-} E[Y|X]$ and $\lim_{X \rightarrow c^+} E[Y|X]$ denote the limiting values of Y when X approaches the cutoff value c from the left and right sides, respectively. Under the continuity assumption in the RD framework that the potential outcome Y is a continuous function of the assignment variable X in absence of the treatment, we have:

$$\lim_{X \rightarrow c^+} E[Y_0|X] - \lim_{X \rightarrow c^-} E[Y_0|X] = 0 \quad (\text{A3})$$

Based on Eqs. (A1)–(A3), if the treatment variable D does jump discontinuously at the cutoff point, we can uncover the treatment effect τ as:

$$\tau = \frac{\lim_{X \rightarrow c^+} E[Y|X] - \lim_{X \rightarrow c^-} E[Y|X]}{\lim_{X \rightarrow c^+} E[D|X] - \lim_{X \rightarrow c^-} E[D|X]} \quad (\text{A4})$$

The above estimator is analogous to the well-known Wald estimator in the IV framework. More generally, it can be shown that under the heterogeneity setting (i.e., $Y_{11} - Y_{j0} = \tau_{ij}$), Eq. (A4) represents the local average treatment effect (LATE) for the subpopulation of compliers:

$$\tau = \lim_{\varepsilon \rightarrow 0} E[Y_1 - Y_0 | D(X = c + \varepsilon) - D(X = c - \varepsilon) = 1] \quad (\text{A5})$$

One direct way to estimate the LATE is to run the one-sided kernel regressions at the cutoff point in X and obtain the point-wise prediction for the four limiting values in Eq. (A4):

$$\lim_{X \rightarrow c^+} E[Z|X] = \frac{\sum_{i: X_i > c} Z_i \cdot K((X_i - x)/h)}{\sum_{i: X_i > c} K((X_i - x)/h)} \quad (\text{A6})$$

$$\lim_{X \rightarrow c^-} E[Z|X] = \frac{\sum_{i: X_i < c} Z_i \cdot K((X_i - x)/h)}{\sum_{i: X_i < c} K((X_i - x)/h)} \quad (\text{A7})$$

where Z can be replaced with Y or D , $K(\cdot)$ is the kernel function, and h is the bandwidth for the kernel regressions. Eqs. (A6) and (A7) are the one-sided point estimates for Y and D based on the observations to the left (or right) of the cutoff point.

In obtaining the consistent estimates of the treatment effect, RD regressions do not require the inclusion of baseline covariates under the no-manipulation assumption that these covariates are continuous on the treatment variable across the cutoff point. In Section 5, we test the validity of this assumption, and also incorporate the covariates into our RD estimation for the robustness check purpose. Following the Erisch–Waugh–Lovell theorem, the covariates can be incorporated into the non-parametric regressions by “residualizing” the outcome variable Y as well as the treatment variable D , i.e., subtracting from Y (or D) its fitted value based on a linear regression of Y (or D) on the covariates.

As mentioned in Section 3, in certain cases the non-parametric estimate of the treatment effect can be numerically equivalent to an IV estimate. To illustrate this point, we formulate the parametric estimation using the following simultaneous equations:

$$Y = \theta + \tau D + f(X - c) + \varepsilon \quad (\text{A8})$$

$$D = \gamma + \delta T + g(X - c) + u \quad (\text{A9})$$

where Eqs. (A8) and (A9) are the second and first stage regressions in the 2SLS framework, θ and γ are the constant terms, and ε and u are the error terms. The binary indicator (T) of being in the treatment group is treated as IV (as it is randomly assigned and also have a strong impact on the probability of being a single child), and τ is the treatment effect to be estimated. In sharp RD designs we have $\delta = 1$, while in fuzzy RD designs δ is the proportion of compliers in the treatment group that falls between 0 and 1. The functions $f(\cdot)$ and $g(\cdot)$ represent the impacts of the assignment variable in the two-stage regressions, and they are usually in polynomial forms to avoid the multicollinearity between X and T (we use the quadratic functions).

Since T is a dummy variable, the 2SLS estimate for τ can be expressed as:

$$\tau = \frac{\sum_{i=1}^N (T_i - \bar{T})(Y_i - \bar{Y})}{\sum_{i=1}^N (T_i - \bar{T})(X_i - \bar{X})} = \frac{E[Y|c < X \leq c + h] - E[Y|c - h \leq X \leq c]}{E[D|c < X \leq c + h] - E[D|c - h \leq X \leq c]} \quad (\text{A10})$$

A comparison between Eqs. (A4) and (A10) indicates that the parametric IV/2SLS estimator for the treatment effect is numerically equivalent to the non-parametric estimator at the cutoff point *if and only if* the uniform kernels are used to estimate the 4 limits in Eq. (A4). In this sense, the IV/2SLS estimator is just a special form of the non-parametric estimator in the RD framework.

References

- Almond, D., Li, H., Zhang, S., 2013. Land Reform and Sex Selection in China. NBER Working Paper No. 19153. <http://www.nber.org/papers/w19153>.
- Angrist, J.D., Krueger, A.B., 1991. Does compulsory school attendance affect schooling and earnings? *Q. J. Econ.* 106 (4), 979–1014.
- Angrist, J.D., Lavy, V., 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Q. J. Econ.* 114 (2), 553–575.
- Angrist, J.D., Lavy, V., Schlosser, A., 2010. Multiple experiments for the causal link between the quantity and quality of children. *J. Labor Econ.* 28 (4), 773–824.
- Banister, J., 1987. *China's Changing Population*. Stanford University Press, Stanford, CA.
- Becker, G.S., 1960. An economic analysis of fertility. *Demographic and Economic Change in Developed Countries*. Princeton University Press.
- Becker, G.S., Lewis, H.G., 1973. On the interaction between the quantity and quality of children. *J. Political Econ.* 81 (2), S279–S288.
- Becker, G.S., Tomes, N., 1976. Child endowments, and the quantity and quality of children. *J. Political Econ.* 84 (4), S143–S162.
- Behrman, J.R., 1996. The impact of health and nutrition on education. *World Bank Res. Obs.* 11 (1), 23–37.
- Black, S.E., Devereux, P.J., Salvanes, K.G., 2005. The more the merrier? The effect of family size and birth order on children's education. *Q. J. Econ.* 120 (2), 669–700.
- Buckles, K.S., Hungerman, D.M., 2013. Season of birth and later outcomes: old questions, new answers. *Rev. Econ. Stat.* 95 (3), 711–724.
- Cameron, L., Erkal, N., Gangadharan, L., Meng, X., 2013. Little emperors: behavioral impacts of China's one-child policy. *Science* 340 (6130), 272–273.
- Chen, Y., Ebenstein, A., Greenstone, M., Li, H., 2013. Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai river policy. *Proc. Natl Acad. Sci.* 110 (32), 12936–12941.
- Cooney, R.S., Wei, J., Powers, M.G., 1991. The One Child Certificate in Hebei Province, China: Acceptance and Consequence, 1979–1988. *Popul. Res. Policy Rev.* 10 (2), 137–155.
- Fan, J., Gijbels, I., 1996. *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London, New York and Melbourne.
- Glewwe, P., Jacoby, H.G., King, E.M., 2001. Early childhood nutrition and academic achievement: a longitudinal analysis. *J. Public Econ.* 81 (3), 345–368.
- Glick, P.J., Marini, A., Sahn, D.E., 2005. Estimating the Consequences of Changes in Fertility on Child Health and Education in Romania: An Analysis Using Twins Data. Cornell University Food and Nutrition Policy Program Working Paper No. 183.
- Guo, Z., Zhang, E., Gu, B., Wang, F., 2003. Diversity of China's fertility policy by policy fertility. *Popul. Res.* 27 (5), 1–10 (In Chinese.).
- Hahn, J., Todd, P., Van der Klaauw, W., 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69 (1), 201–209.
- Hatton, T.J., Martin, R.M., 2010. Fertility decline and the heights of children in Britain, 1886–1938. *Explor. Econ. History* 47 (4), 505–519.
- Hesketh, T., Lu, L., Xing, Z.W., 2005. The effect of China's one-child family policy after 25 years. *N. Engl. J. Med.* 353 (11), 1171–1176.
- Huang, S., Sun, T., 2013. The balance between burden of population and demographic dividend. *Acad. Mon.* 45 (7), 66–77 (In Chinese.).
- Imbens, G., 2010. Better LATE than nothing: some comments on Deaton (2009) and Heckman and Urzua (2009). *J. Econ. Lit.* 48 (2), 399–423.
- Imbens, G., Kalyanaraman, K., 2012. Optimal bandwidth choice for the regression discontinuity estimator. *Rev. Econ. Stud.* 79 (3), 933–959.
- Imbens, G., Lemieux, T., 2008. Regression discontinuity designs: a guide to practice. *J. Econom.* 142 (2), 615–635.
- Lee, D.S., 2008. Randomized experiments from non-random selection in U.S. house elections. *J. Econom.* 142 (2), 675–697.
- Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. *J. Econ. Lit.* 48 (2), 281–355.
- Lei, X., Tan, L., Zhao, Y., 2010. Does retirement affect health? *China Econ. Q.* 9 (4), 1539–1558 (In Chinese.).

- Li, H., Zhang, J., Zhu, Y., 2008. The quantity-quality trade-off of children in a developing country: identification using Chinese twins. *Demography*. 45 (1), 223–243.
- Li, J., 1995. China's one-child policy: how and how well has it worked? A case study of Hebei province, 1979–1988. *Popul. Dev. Rev.* 21 (3), 563–585.
- Li, J., Cooney, R.S., 1993. Son Preference and the One Child Policy in China: 1979–1988. *Popul. Res. Policy Rev.* 12 (3), 277–296.
- Li, S., Xing, C., 2010. China's Higher Education Expansion and its Labor Market Consequences. IZA Discussion Paper No. 4974.
- Liu, H., 2014. The quality-quantity trade-off: evidence from the relaxation of China's one-child policy. *J. Popul. Econ.* 27 (2), 565–602.
- Liu, Y., Lu, M., 2008. How could loosening the one-child policy affect China's economic growth? A theoretical analysis based on within-household old-age support. *China Econ. Q.* 7 (4), 1271–1300 (In Chinese.).
- Lordan, G., Frijters, P., 2013. Unplanned pregnancy and the impact on sibling health outcomes. *Health Econ.* 22 (8), 903–914.
- Ma, X., 2010. Discuss China's future birth control policy from a perspective of population aging: empirical analysis on the causes. *Econ. Vis.* 2010 (05), 46–49 (In Chinese.).
- McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: a density test. *J. Econom.* 142 (2), 698–714.
- Nichols, A., 2007. Causal inference with observational data. *Stata J.* 7 (4), 507–541.
- Park, C., Chung, W., 2012. Sib-ship size, birth order, and children's education in developing countries: evidence from Bangladesh, Hitotsubashi. *J. Econ.* 53 (1), 1–23.
- Qian, N., 2009. Quantity-Quality and The One Child Policy: the Only-Child Disadvantage in School Enrollment in Rural China. NBER Working Paper No. 14973. <http://www.nber.org/papers/w14973>.
- Rosenzweig, M.R., Zhang, J., 2009. Do population control policies induce more human capital? Twins, birth weight and China's 'one-child' policy. *Rev. Econ. Stud.* 76 (3), 1149–1174.
- Smith, J.P., 2004. Unraveling the SES-health connection. *Popul. Devel. Rev.* 30, 108–132.
- Thistlethwaite, D.L., Campbell, D.T., 1960. Regression-discontinuity analysis: an alternative to the ex post facto experiment. *J. Educ. Psychol.* 51 (6), 309–317.
- Van der Klaauw, W., 2002. Estimating the effect of financial aid offers on college enrollment: a regression-discontinuity approach. *Int. Econ. Rev.* 43 (4), 1249–1287.
- Willis, R.J., 1973. A new approach to the economic theory of fertility behavior. *J. Political Econ.* 81 (2), S14–S64.
- Xu, J., 2010. Analysis of Chinese Population Policies' Effect (MA thesis). Jilin University, Jilin.
- Yang, F., 2003. Historical Research on Family Planning of Contemporary China (Ph.D. diss.). Zhejiang University, Zhejiang (In Chinese.).
- Zhang, J., Spencer, B.G., 1992. Who signs China's one-child certificate, and why? *J. Popul. Econ.* 5 (3), 203–215.
- Zhou, J., Yin, Y., 2011. The effect of China's birth control policy on household saving rate based on provincial level panel data. *J. Financial Res.* 2011 (10), 61–73 (In Chinese.).