One-Child Policy and the Rise of Man-Made Twins*

Wei Huang, Xiaoyan Lei, and Yaohui Zhao

Abstract: This paper examines an unintended response to the One-Child Policy in China: twinning births. Analysis of population census data shows that the One-Child Policy has accounted for more than one-third of the increase in the reported births of twins since the 1970s. Investigation using birth space with prior births and height difference within twins suggests that the increase in births of twins is partly due to parents reporting regularly-spaced children as twins to avoid the policy violation punishment. The study highlights the possibility of individual behavioral response to undesirable government policies and the potential social consequences.

JEL Codes: J08, J11, J13; Keywords: Twins, One-Child Policy, China

---

* Huang: Department of Economics, Harvard University (email: weihuang@fas.harvard.edu). Lei: National School of Development, Peking University (email: xylei@nsd.pku.edu.cn). Zhao: National School of Development, Peking University (email: yhzhao@nsd.pku.edu.cn). We thank Gary Becker, David Cutler, Avraham Ebenstein, Richard Freeman, Edward Glaeser, Claudia Goldin, James Heckman, Adriana Lleras-Muney, Lawrence Katz, Mark Rosenzweig, John Strauss, T. Paul Schultz, David Wise as well as the editor and anonymous referees for their suggestions. We also thank the participants at PAA conference, China Meeting of the Econometric Society, CGEB/LAEF Conference on Economics and Demography and China Economics Summer Institute for their comments. Wei Huang also thanks the financial support from the Cheng Yan Family Research Grant at Harvard, Graduate Student Research Grant at IQSS and NBER Pre-doc Fellowship. All errors are ours. Contact Xiaoyan Lei if any comments.

"One is too few," said a woman (in China) waiting at the hospital. "People want to have a second child." —Elizabeth Grether, ABC News, August 3, 2011

1. Introduction

Since Rosenzweig and Wolpin (1980a, 1980b), an established strand of the economics literature has used samples of twins to deal with biases from unobserved factors, particularly in estimating educational returns and the effects of family size (e.g., Ashenfelter and Krueger 1994). This methodology has been further developed and followed by studies in various contexts, including developing countries such as China (e.g., Li et al. 2008; Rosenzweig and Zhang 2009).

Previous studies generally assume that twins are born randomly across the population conditional on observed biological factors, such as the mother's age at childbirth and ethnicity. But the studies do not consider the possibility that the birth of twins, as a fertility behavior, could be manipulated in response to the distorted incentives imposed by relevant policies. For the first time, our paper investigates the impact of fertility policy on births of twins to examine the behavioral responses to the distorted incentives of the One-Child Policy (OCP) in China.

A first glance, the timing of OCP and the trend in the rate of twin births suggest strong correlation between them (Figure 1): the rate of twin births reported in population censuses more than doubled between the late 1960s and the early 2000s, from 3.5 to 7.5 per thousand births. The rate of increase in twin births was greater after 1979, when OCP was fully implemented.[1]

[1] The pattern of increasing twins in China is not unique and the rate of twin births is

[Figure 1 about here]

There are reasons to believe that the correlation may not be accidental. When a couple is allowed one birth, the only legitimate way to have two children is to give birth to twins, which is supposedly out of the control of the couple. One option for achieving this goal is to take fertility drugs. Although fertility drugs are meant to be used for infertility by inducing ovulation, there is some anecdotal evidence that women sometimes intentionally take fertility drugs to obtain twins.[2] For those who fail to have twins, an alternative option is to report fake twins, i.e., to register two consecutive siblings as twins. For example, officials in Yunnan province identified 700 pairs of fake twins in 342 villages in 2000. More surprisingly, among 23 pairs of twins reportedly born in 1999, 18 were identified as fake.[3]

Using the timing and geographical variation of policy violation fines, we find that the policy accounts for at least one-third of the increase in twin births since the 1970s. Such

---

heterogeneous across different populations. For example, the rate of twin births increased from 18.9 to 33.3 per thousand births from 1980 to 2009 in the United States (Martin et al. 2012). Therefore, no conclusion can be drawn before more serious analyses.

[2] Because of the lack of stringent regulation, these drugs are easily accessible, for example through online pharmacies and private hospitals. News Source: http://abcnews.go.com/Health/chinese-women-fertility-drugs-bypass-child-policy/story?id=14219173.

[3] News Source (Chinese): http://www.people.com.cn/GB/channel1/13/20000728/163617.html.

policy-associated twins are more likely to be found in rural areas and in *observed* second births.[4] The results are robust to a set of alternative regression specifications.

After getting evidence consistent with the hypothesis of man-made twins, we further investigate the mechanism through which this happens, i.e., whether they get twins by measures like reporting fake twins. We find that the birth gap between the first birth and the *observed* second twin birth is 0.08 year longer relative to an *observed* second single delivery after the policy was introduced. We also find that height difference within twins is larger where the OCP fine rate is higher. These findings are consistent with the hypothesis that OCP incentivizes parents to report non-twin children as twins.

The structure of the paper is as follows. Section 2 provides background on China's OCP. Section 3 describes the data and the empirical results. Section 4 explores the potential measures that people take to have twins. Section 5 concludes with a discussion of the findings, policy implications, and suggestions for further research.

2. Background

OCP was introduced in 1979 to alleviate social, economic, and environmental problems in China (Greenhalgh 1986; Greenhalgh and Bongaarts 1992; Wang 2012). Legal measures, such as monetary penalties and subsidies, have ensured the effective enforcement of OCP since 1979.

[4] Note that *observed* second births may not be *real* second births if the parents report fake twins. As discussed later, the parents possibly report the second birth child and the third birth child together as twins, and these twins will be observed as the second birth.

Because of the heterogeneous regional development across China, Central Party Committee "Document 7" devolved responsibility from the central government to the local and provincial governments. The devolution allowed for regional variation in family planning policies, such as the amounts for monetary penalties or subsidies (Greenhalgh 1986). However, OCP mainly focused on the Han ethnicity, the largest ethnic group in China, with more than 90 percent of the population.

In addition to the timing of its implementation, the additional measure of OCP in this study is the average monetary penalty rate for one unauthorized birth in the province-year panel from 1979 to 2000.[5] The OCP regulatory fine (policy fine) is formulated in multiples of annual income (Ebenstein 2010; Wei and Zhang 2011).[6]

Since a period of approximately nine months is needed from the beginning of pregnancy until birth, parents' decision to have a child should be made close to a year in advance. For each birth, we construct a variable, *policy fine rate*, which is the weighted mean value of the fine rate in the 12 months just before the pregnancy in a given province.[7] The effective fine rate was zero

[5] Details on the construction of this variable can be found in Ebenstein (2010).

[6] Appendix A provides the details.

[7] Because the 1982 Census does not have birth month information, we assume the children surveyed in 1982 were born in June and conduct the same procedure. The estimates do not rely on the OCP measure we constructed here: results are consistent if we simply use the fine rate one year before the child was born. (The results are available upon request.) When matching the policy fine rate to the current local province, we assume that the province of birth is the province

for children born before 1979, when OCP started. We drop the children born after 2001, because the fine rate is not available after 2000. Appendix Table A1 presents the summary statistics: the rate of twin births is 0.58 percent for Han (the majority) and 0.44 percent for minorities.

3. Data and Empirical Results

3.1 Data

The main data used in this study are from the 1982, 1990, and 2000 Population Censuses and the 2005 One-Percent Population Survey. A detailed description of the data and sample restrictions is provided in Appendix A. Twins are defined as children in the same household with the same birth year and birth month.[8] Observations are at the birth level, so twins are treated as a single observation, because they are in the same observed birth.

3.2 Impact of OCP on Twinning Rates in China

To evaluate the effects of OCP on the twinning rate, we estimate the following equation:

$$(1)\ Twin_{ijky} = \beta_0 + \beta Fine_{jym} + \delta_k + \delta_y + \delta_{ky} + \delta_j + X_{ij} + \varepsilon_{ij}$$

---

of current residency, which may not be true due to migration. Using census data after 1990 with information on place of birth, we found over 95 percent of the children live in the same province where they were born, indicating that interprovincial migration should not be a big issue of concern in the analysis.

[8] Because the 1982 Census data do not have information on birth month, we define twins in that year as those children born in the same household with the same birth year only. The results are almost the same when we drop the 1982 Census or define twins only using the year of birth in all the other data sets.

where the dependent variable, $Twin_{ijky}$, denotes whether birth i in year y and province j is a twin birth in survey year k. $Fine_{jym}$ is the OCP fine rate defined above in province j for children born in year y and month m. The main coefficient of interest, β, gives the association of the OCP fine with the reported twinning rate and is interpreted as the impact of OCP.

δ_k, δ_y and δ_{ky} are indicators for year of birth y, survey year k, and their combinations, respectively. δ_j denote the province dummies. X_{ij} is a set of covariates, including dummies for residence type (urban/rural), parents' ethnicity (both Han or either a minority), birth order, birth month, mother's education level, and mother's age at childbirth as well as the provincial specific linear trends in birth cohorts.

The first three columns in Table 1 report the OLS estimates for β in Equation (1) with standard errors clustered at the provincial level. The results indicate that an increase equivalent to one year's income in the policy fine is associated with a 0.066 percentage point increase in the twin birthrate among the whole sample. The estimates in column 2 suggest that 36 percent of the increase in twins in the Han ethnicity sample can be attributed to OCP.[9] As expected, the association and significance survive in the sample for Han ethnicity but diminish in the sample for minority groups.

[Table 1 about here]

[9] The twinning rates before and after OCP are 0.39 percent and 0.67 percent, respectively. The mean value of the fine rate increases from 0 to 1.4 years of local household income. The part of the increase in the twinning rate that can be explained is 0.072*1.4/0.28 = 0.36.

We further interact the policy fine with Han ethnicity and report the results in column 4. Consistent with the above, the main effect of the fine diminishes and the interaction is positive and significant, with similar magnitude as in column 2.

Li et al. (2011) argue that spatial and temporal variation of the OCP policy may be endogenous. They find that the policy fine increases with community wealth and the local government's birth-control incentives and decreases with the local government's revenue incentives. We test the endogeneity of the fine rate in Appendix B and find no evidence of endogeneity under the settings of this paper. Therefore, we conduct another set of regressions without using the policy fine rate, exploring only the timing of OCP:

$$(2)\ Twins_{ijky} = \beta_0 + \beta_2(Policy_{y\geq 1980} \times Han_i) + \delta_k + \delta_y + \delta_{ky} + \delta_j + X_i + \varepsilon_i$$

where $Policy_{t\geq 1980}$ denotes an indicator of whether birth *i* was in 1980 or after, and Han_i is an indicator for Han ethnicity of both parents. This difference-in-differences (DID) estimate in the final column of Table 1 indicates that OCP explains over 54 percent of the increase in twins, which is larger than that from estimates based on the policy fine rate.[10] The result is reasonable, since the policy fine captures only one means of punishment and the fine rate is averaged at the provincial level, which may miss some contributing variation within a province.

An important assumption of the DID estimation is that the trend for Han ethnicity without the policy would be similar to that for the minority group. Figure 2 examines this assumption by plotting the twin birthrate by the ethnicity group of the parents against the year of birth. As the

[10] The proportion that can be explained by OCP is 0.15/0.28 = 0.54.

figure shows, the trends in the two groups are almost identical prior to the introduction of OCP, indicating that the two trends without the policy are likely to be the same.

[Figure 2 about here]

3.3 Heterogeneous Impact of OCP on Twinning Births, by Type of Residence and Birth Order

Enforcement of OCP differs in urban and rural areas. For example, urban areas strictly enforce the policy, while many rural areas allow a couple to have a second child if the first is a girl. This varying enforcement, together with other potential differences between the two areas, may result in heterogeneous effects.

The regression results by residence are reported in Table 2. The policy fine is positively correlated with the incidence of twins in urban and rural areas. The association in rural areas is larger and more significant, indicating the incentive to have twins in rural areas may dominate that in urban areas.

[Table 2 about here]

The incentive to have a twin birth may be stronger at different points in the two areas: urban parents may have to manipulate the birth date for the first child, while rural parents can wait until the second. We examine this heterogeneity by interacting policy fines with birth order dummies in the regressions. As shown in panel B of Table 2, twins in the second birth are the most policy relevant and the association is mainly reflected in rural areas, which is consistent with OCP enforcement.[11]

[11] Note again that the second birth here may not be the real second birth, because fake twins may

4. *Mechanisms: Reporting Fake Twins and/or Taking Fertility Drugs*

As mentioned above, people may either report single children as twins *ex post* or take fertility drugs *ex ante* to raise the probability of multiple children in a single birth. In this section, we try to identify the two channels by examining the birth gap between the first two *observed* births and the height difference within twins, respectively.

4.1 Impact of OCP on Birth Gap between the First Two Observed Births

If an elder child is registered with a younger one as twins, the birthdate of the reported twins tends to be registered as that of the younger child, since the parents have to wait until both children are born. If fake twins are reported to be born as *observed* second births, there should be a longer birth gap between the first two *observed* births because it is actually the gap between the first birth and the *actual* third birth. If parents plan to have twins by taking fertility drugs, however, there is no reason why they would report a delivery date that is later than normal. Therefore, we restrict the sample to the *observed* second births and conduct the following regressions:[12]

---

be reported. As discussed in section 4, parents might report the second and third births together as twins, who are misleadingly seen as second births.

[12] It is possible that parents have twins in the first *observed* birth. We do not consider this case here, because it is not possible to calculate the gap between the first observed birth and the previous one, and the results in the previous section suggest that policy-related twinning is mostly concentrated in observed second births.

$$(3)\ Birth\ Gap_i = \beta_0 + \beta_3\left(Policy_{y\geq 1980}\times Twins_i\right) + \beta_3' Twins_i + \delta_k + \delta_y + +\delta_{ky} + \delta_j + \mathrm{X}_i' + \varepsilon_i$$

In Equation (3), the variable $Birth\ Gap_i$ denotes the *observed* birth gap between the current (second) and previous (first) delivery for birth i. $Twins_i$ denotes whether the current birth of the same parents for birth i is a twin birth or not, which captures the potential difference, if any, in the birth gap between single births and twin births driven by factors other than OCP. The coefficient, β_3, on the interaction term is of central interest because it reflects how much additional time is needed to give birth to twins than to a single child after OCP was implemented. If there are reported fake twins in the *observed* second births after OCP, we should expect the estimated coefficients for β_3 to be positive, because the birth gap between the first and second *observed* births would be lengthened by a report of fake twins.

The covariates $\delta_k, \delta_y, \delta_{ky}$ and δ_j have the same definitions as before. X_i' includes dummies for residence type parents' ethnicity, mother's education level, and mother's age at first childbirth as well as the provincial specific linear trends in birth cohorts.

Table 3 reports the OLS estimates for β_3 and β_3' for different samples. For the full sample (column 1), we obtain a positive and significant estimate for β_3, showing that a twin birth needs an additional 0.08 year to achieve than a singleton birth does after the OCP was introduced. The next two columns provide results for Han ethnicity and minorities, respectively, and the significant estimates only appear in Han ethnicity.

[Table 3 about here]

4.2 Impact of OCP on Height Difference within Twins

In this section, we use height differences of the two children within a twin birth to test whether

they are "man-made". We do the test for same-gender twins and mixed-gender twins, as the results of each have different implications. The detailed derivations are relegated to Appendix C due to space constraints. Here we summarize some basic results from the derivation: (*a*) if OCP is not relevant for twin births, the height difference within twins should be *uncorrelated* with the policy; (*b*) if OCP leads to more fertility drug use rather than reporting fake twins, a larger height difference is predicated to exist *only* within same-gender twins under OCP; and (*c*) if there are any reported fake twins, a larger height difference is expected *both* within same-gender twins and within different-gender twins.

For this analysis, we turn to the data from the China Health and Nutrition Survey (CHNS), which provides information on height. We try to use the same definition of twins as that with the Census data in the above analyses; a detailed description of the data and sample restrictions is in Appendix A. To estimate the relationship between height difference and OCP, we conduct the following estimation:

$$(4)\ HD_{ijyk} = \alpha + \beta_4 Fine_{jym} + \delta_y + \delta_j + \delta_k + Z_i + e_i$$

where the dependent variable, HD_{ijyk} denotes the height difference within twin pair i in province j born in year y of wave k. The coefficient β_4 gives the association between the policy fine and the height difference within twins. Because there are only 72 pairs of twins observed in total and 33 different pairs in the regressions, we just combine the birth years into four groups (every five years as one group), δ_y, to capture the birth year effects. We also combine the neighboring provinces into five groups and control for the region fixed effects δ_j. The other covariates, Z_i, include indicators for whether the twins are same-gender twins, urban residence,

whether the boy within different-gender twins is taller than the girl, and continuous variables like average height of pair *i*, mother's age at childbirth, and age and age squared of the twins.

Panel A in Table 4 reports the estimates of Equation (4) with standard errors clustered at the provincial level. The first column shows that increasing the fine by one year's income is associated with an increase in the height gap of twins by 1.8 centimeters, suggesting that couples may have employed methods to make twins. Interacting the fine variable with the indicator variable of same- and different-gender twins in column 2 shows that the policy fine is positively associated with height differences for both types of twins. Columns 3 and 4 use the gap/mean ratio as the dependent variable and the results are similar.

[Table 4 about here]

Since CHNS is a panel data set and the same twins may be surveyed in different waves, we alternatively keep observations from only the latest wave for each pair and run the same regressions as above. Panel B in Table 4 reports the estimates for the key variables and the results represent the same pattern, with coefficients of larger magnitude than those in panel A. Altogether the results in this panel provide supportive evidence of parents reporting consecutive children as twins to avoid the policy violation punishment.

5. Conclusions and Discussion

In 1979, the Chinese government launched the One-Child Policy, which led to hundreds of millions of couples involved in this strict family-planning program lasting for more than 30 years. In this paper, we find that an increase in the policy fine of one year's income is associated with an increase in twin births by approximately 0.07 per thousand births, indicating that at least

one-third of the increase in twins since the 1970s can be explained by OCP. We then examine the heterogeneous effects by residence and find that the impact of the policy is larger in rural areas, where the policy raises the twin birthrate mainly for second births.

Furthermore, we find that since OCP was put into effect, the birth gap between the first two *observed* births is 0.08 year longer when the *observed* second birth is a twin than when it is not. In addition, the height difference within twins is positively associated with the policy fine and the association exists in both same- and different-gender twins. These findings support the hypothesis that OCP has incentivized people to have twins by reporting non-twin children as twins.

Since behavior response is related to deadweight loss of social welfare (Hendren 2013), economics literature usually examines individuals' behavioral response to or against government policies. This study builds up the literature by examining people's behavioral response to OCP. Our estimates indicate a sizeable behavioral response from Chinese couples. This finding should motivate future studies to examine individuals' behavioral responses to other undesirable public policies, and calculate the associated welfare gain or loss, especially in China, the largest developing country. This study also helps to explain the context and background of studies that use data on twins in China. Since couples can intentionally have twin births to bypass OCP, the distribution of reported twins in China may not be random. It is worth noting that the results do not rule out the possibility that women may take fertility drugs to have twins, and that the results allude to the importance of carefully screening twins with observable characteristics when analyzing Chinese data sets of twins.

There are some possible limitations in the current study and we leave these issues for future study. First, the evidence for fake twins provided in this paper is suggestive rather than determinant, because we cannot observe fake twins directly in our data. Identification of twins would require careful and detailed field surveys. It would be interesting to see how many twins are fake and how these fake twins are distributed across household socioeconomic status. Second, although CHNS is the largest data set we can find for the analysis on height difference, the sample size is too small; therefore, the results are sensitive to model specification and any attempt to generalize the results should be treated with caution. We hope the issue can be better addressed in the future with a larger data set. Finally, the fine rate measure we use in this study captures only one dimension of OCP. We hope that future studies can explore more comprehensive measures of the policy to test our story and investigate which dimensions of the policy lead to more fake twins.

References

Ashenfelter, O., and A. Krueger. 1994. "Estimates of the Economic Return to Schooling from a New Sample of Twins." *The American Economic Review* 84(5): 1157–73.

Ebenstein, A. 2010. "The 'Missing Girls' of China and the Unintended Consequences of the One Child Policy." *Journal of Human Resources* 45(1): 87–115.

Greenhalgh, S. 1986. "Shifts in China's Population Policy, 1984–1986: Views from the Central, Provincial, and Local Levels." *Population and Development Review* 12(3): 493–515.

Greenhalgh , S., and J. Bongaarts. 1992. *Fertility Policy in China: Future Options* (pp. 401–19). Springer US.

Hendren, N. 2013. "The Policy Elasticity." NBER Working Paper No. 19177, National Bureau of Economic Research, Cambridge, Mass.

Li, H., J. Yi, and J. Zhang. 2011. "Estimating the Effect of the One-Child Policy on the Sex Ratio Imbalance in China: Identification Based on the Difference-in-Differences." *Demography* 48(4): 1535–57.

Li, H., J. Zhang, and Y. Zhu. 2008. "The Quantity-Quality Trade-off of Children in a Developing Country: Identification Using Chinese Twins." *Demography* 45(1): 223–43.

Martin, J. A., B. E. Hamilton, S. J. Ventura, M. J. Osterman, E. C. Wilson, & T. J. Mathews. 2012. Births: final data for 2010. *National vital statistics reports*, *61*(1), 1-72.

Rosenzweig, M. R., and K. I. Wolpin. 1980a. "Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models." *Journal of Political Economy* 88(2): 328–48.

———. 1980b. “Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment.” *Econometrica: Journal of the Econometric Society* 48(1): 227.

Rosenzweig, M. R., and J. Zhang. 2009. “Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China’s ‘One-Child’ Policy.” *The Review of Economic Studies* 76(3): 1149–74.

Wang, F. 2012. “Family Planning Policy in China: Measurement and Impact on Fertility.” Working Paper

Wei, S. J., and Zhang, X. 2011. The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China. *Journal of Political Economy* 119(3): 511–64.

Appendix

Appendix A: Data

A1. Census Data

The data used in this study are from the 1982, 1990, and 2000 Population Censuses and the 2005 One-Percent Population Survey. All the data sets contain years of birth, region of residence, type of residence (urban/rural),[13] gender, ethnicity, education, and relation to the household head. The data sets after 1982 also include month of birth. For women older than 15, the data also provide information about their fertility history, including number of children ever born and number of living children.

For the analysis of twin births and family background, we first keep only those households with at least one child and with information available for the mother. We restrict the sample to those whose household heads and spouses are their first marriage and further restrict the sample to those households with equal numbers of reported living children, children ever born, and children observed in the survey. Doing so ensures that all the children in each household have

[13] The 1990 Census does not provide the type of current residence; it provides information on whether the respondents lived in the same place five years ago and what the type of residence was then. Therefore, we combine information from these two variables and construct an indicator for type of current residence. More specifically, we keep the respondents who lived in the same place as five years ago; then we calculate for each area the proportion of urban and rural, respectively, five years ago. The residence type with the higher proportion was used as the proxy for current residence type for all the observations in this area.

the same mother observed in the household and that the sample covers the information needed for the children. In case we miss children who have moved from the household, we further drop households with children over age 17 in the survey. We finally drop households where the mother's age at childbirth is either younger than 15 or older than 50, as these subsamples may be too special or may contain recording errors.[14]

Appendix Table A1 presents the means and standard deviations for the key variables.

[Appendix Table A1 about here]

A2. CHNS

The China Health and Nutrition Survey (CHNS) includes 26,000 individuals in nine provinces that contain approximately 56 percent of the population of Mainland China. The nine provinces vary substantially in geography, economic development, public resources, and health indicators. Data collection began in 1989 and has been implemented every two to four years since then (Jones-Smith and Popkin 2010).

We first keep the children younger than 18 years in CHNS and select out the twins. Twins are defined as the children with exactly the same birth year and birth month within the same household. We define the height difference of each pair as the difference between the taller and the shorter child. Considering that the height gap may change as the children grow up, we also introduce another measure, the ratio of the height gap to the mean height of the pair (gap/mean). We match the twin sample with the fine data and drop those born after 2001, as we did for the

[14] We also drop data for Tibet.

Census sample. After these restrictions, we have 72 pairs in total, among whom 53 are same-gender twins and the rest are different-gender twins. Appendix Table 2 reports the summary statistics for the height gap and the gap/mean ratio. The mean height gap is 2.2 centimeters and the gap/mean ratio is 1.75 percent. As expected, different-gender twins have a much larger difference in height than same-gender twins (4.65 versus 1.30 cm).

[Appendix Table A2 about here]

A3 One-Child Policy Regulatory Fine

Appendix Figure A1 shows the pattern of the policy fine in 1980–2000 in each province. The figure shows that the fines in the different provinces generally follow different patterns, in timing and magnitude. For example, Liaoning province raised the fine from one year's income to five in 1992, while Guizhou raised the fine from two to five years of income in 1998 and Hunan from one to two years of income in 1989. The geographical and temporal variation helps us identify the impact of OCP on the reported births of twins in the following empirical analysis.

[Appendix Figure A1 about here]

Appendix B: Endogeneity Tests of the OCP Fine Rates

Li et al. (2011) argue that spatial and temporal variation of OCP may be endogenous. They find that the policy fine increases with community wealth and the local government's birth-control incentives and decreases with the local government's revenue incentives. In particular, spatial and temporal variation in OCP may be affected by the local fertility rate, which, in turn, may correlate with the incidence of twins. To test this possibility, we regress future policy fines on

prior twin birthrates and see if the latter has predictive power on the former. If the association is significant, then the endogeneity problem is worthy of concern. More specifically, in each regression, we use whether the observation is a twin birth as the key independent variable and the amount of policy fine required in the next year, three years later, or five years later as the dependent variable, respectively. Because the local government started to have the local policies in 1984, the year of "Document 7", I kept the post-1980 birth cohorts to check this. As shown in Appendix Table B1, the rate of twin births does not seem to have any predictive power on the amount of the policy fine over the next one, three, or five years, suggesting that the reverse causality problem may not be serious in this study. Also note that the coefficients are really small, which suggest that the predicted fine rate would change smaller than 0.01 even if the current twining rate had increased from 0 to 10 percent.

[Appendix Table B1 about here]

Appendix C: The Two Hypotheses and the Derivation of Their Testable Implications

In this part, we derive testable implications for two different hypotheses, the real "man-made" twins hypothesis and the false twins hypothesis. For simplicity, we only use financial penalties (fines) to measure OCP and assume it equals one when a financial penalty policy is established in the local province and zero otherwise.

C1. Real "Man-Made" Twins Hypothesis

The man-made twins hypothesis indicates that individuals are motivated to use technology, such as fertility drugs, to give birth to twins under OCP. Taking fertility drugs should be reasonable,

but embryo technologies did not appear in China until the late 1990s. Fertility drugs are usually used to induce ovulation in women with an infertility problem (Rossing et al. 1994). When a woman takes a fertility drug, the possibility of multiple ovulations and thus the likelihood of having twins is raised (Bortolus et al. 1999; Starr 2008). Fertility drugs are classified as prescription medicines, but some people may purchase them in certain private hospitals or obtain prescriptions from certain doctors in an illegal way, such as bribing the doctors.

Under the man-made twins hypothesis, individuals are more likely to take fertility drugs (Take = 1) to have more children and avoid being punished under OCP, that is, $\Pr(Take = 1|Fine = 1) > \Pr(Take = 1 \,|\, Fine = 0)$. Taking certain fertility drugs or using technologies increases the probability of giving birth to twins: $\Pr(Twins = 1|Take = 1) > \Pr(Twins = 1 \,|\, Take = 0)$. We also assume that, conditional on individuals' behaviors (e.g., Take), giving birth to twins is independent of OCP. In addition, the biological function of fertility drugs or embryo technologies is to develop multiple zygotes in the uterus at the same time, rather than to stimulate the split of a single fertilized egg into two or more embryos. Thus, these actions only raise the probability of dizygotic (DZ) twins rather than that of monozygotic (MZ) twins. That is,

$\Pr(DZ = 1|Take = 1) > \Pr(DZ = 1 \,|\, Take = 0)$ and

$\Pr(MZ = 1|Take = 1) = \Pr(MZ = 1 \,|\, Take = 0)$.

MZ twins are nearly identical and they are always of the same gender, except in the extreme case when there is a mutation during development. Certain characteristics of MZ twins become more alike as twins age, such as IQ and personality (Segal 1999). DZ twins, however, like other

siblings, have an extremely small chance of having the same chromosome profile. DZ twins may look very different from each other and may be of different genders or of the same gender. In some sense, DZ twins can be viewed as normal siblings who happen to be the same age. Therefore, we have $\Pr(DZ = 1|DG = 1) = 1, \Pr(SG = 1 \mid MZ = 1) = 1$, $0 < \Pr(\mathrm{DZ} = 1|\mathrm{SG}) < 1$, and $0 < \Pr(\mathrm{MZ} = 1|\mathrm{SG}) < 1$.

Because of genetic disparity, DZ twins tend to have more differences than MZ ones as they grow up, including differences in height (Fischbein 1977; Smith et al. 1973), weight (Stunkard et al. 1986), mental ability profiles (Segal 1985), bone mass (Smith et al. 1973), and so on. Specifically, Fischbein (1977) found that MZ twins have a significantly higher concordance in height than for DZ pairs during puberty, for both boys and girls, and yearly height increments are also more similar for MZ pairs, indicating that the height spurt occurs more simultaneously for MZ twins in comparison with DZ twins. Therefore, the height difference (HD) within a DZ (same-gender) pair should be larger than that of an MZ pair if other factors are equalized. That is, $E(HD|SG = 1, DZ = 1) > E(HD|SG = 1, MZ = 1)$.

In addition, we assume that the actions people take do not influence twins' height differences conditional on the twins' type (MZ or DZ). Based on the aforementioned facts and assumptions, it can be shown that

(1) E(HD|Fine=1,Twins=1)>E(HD|Fine=0,Twins=1),

(2) E(HD|Fine=1,SG=1)>E(HD|Fine=0,SG=1), and

(3) E(HD|Fine=1,DG=1)=E(HD|Fine=0,DG=1).

Equation (1) states that the height difference within twins should be larger with OCP than

without, as there will be more DZ twins because of the fertility-stimulating activities parents undertake in response to OCP. Further, because these activities raise the proportion of DZ twins in same-gender twins, the height difference in this group will become larger (Equation (2)). The height difference in the different-gender twins will not change, because they are DZ anyway (Equation (3)). These results provide a way to test whether a woman took measures (e.g., fertility drugs) to stimulate the birth of twins, which we do not observe directly in the data.

C2. Reporting False Twins Hypothesis

The false twins hypothesis is that parents report non-twin children as twins to avoid punishment under OCP. This was feasible because of the following special circumstances earlier in China. First, many women gave birth at home in the 1980s and it was easy for parents to hide their births if they intended to do so. Second, birth certificates were not launched until 1997 and children's birthdates were easy to manipulate before that. Third, children—especially siblings—look alike, in particular when the age difference is not large, which made it possible to report them as twins.

Under the false twins hypothesis, OCP stimulates people's incentives to report false twins, that is, $Pr(Twins^*|Fine=1)>Pr(Twins^*|Fine=0)$, in which $Twins^*$ denotes the observed twins, including real ones and false ones. For real twins (Twins), we assume all of them are reported, that is, $Pr(Twins^*|Twins)=1$.

The most important difference between false twins and true twins is the age of the children. Because of the age difference, the height difference within false twins should be larger, so $E(HD|Twins^*)>E(HD|Twins)$ *if* $Pr(Twins|Twins^*)<1$.

The condition $Pr(Twins|Twins^*)<1$ ensures that false twins do exist. If parents have a strong preference for children and do not care about the gender, then the gender composition of false twins should be random, so that height differences in (observed) same-gender twins and different-gender twins should be larger under OCP. However, if parents have a strong boy preference, they are more likely to construct different-gender twins because they have less incentive to make false twins when they had a boy already. No matter which case it is, under the false twins hypothesis, we must have

(4) $E(HD|Twins^*,Fine=1)>E(HD|Twins^*,Fine=0)$ and

(5) $E(HD|DG^*,Fine=1)>E(HD|DG^*,Fine=0)$,

in which DG^* denotes the observed different-gender twins. As before, Equations (4) and (5) are based on observables, so they can be tested with empirical analysis.

In summary, from the derivations above, we reach the following testable results: (*a*) if OCP is not relevant for twin births, the height difference within twins should be *uncorrelated* with the policy; (*b*) if OCP leads to more fertility drug use rather than reporting false twins, we are likely to see a larger height difference *only* in same-gender twins under OCP; and (*c*) if there are any reported false twins, a larger height difference is expected *both* in same-sex and different-sex twins under OCP.

Additional References

Bortolus, R., F. Parazzini, L. Chatenoud, G. Benzi, M. M. Bianchi, and A. Marini. 1999. "The Epidemiology of Multiple Births." *Human Reproduction Update* 5(2): 179–87.

Fischbein, S. 1977. "Intra-Pair Similarity in Physical Growth of Monozygotic and of Dizygotic Twins during Puberty." *Annals of Human Biology* 4(5): 417–30.

Jones-Smith, J. C., and B. M. Popkin. 2010. "Understanding Community Context and Adult Health Changes in China: Development of an Urbanicity Scale." *Social Science & Medicine* 71(8): 1436–46.

Li, H., J. Yi, and J. Zhang. 2011. "Estimating the Effect of the One-Child Policy on the Sex Ratio Imbalance in China: Identification Based on the Difference-in-Differences." *Demography* 48(4): 1535–57.

Rossing, M. A., J. R. Daling, N. S. Weiss, D. E. Moore, and S. G. Self. 1994. "Ovarian Tumors in a Cohort of Infertile Women." *New England Journal of Medicine* 331(12): 771–76.

Segal, N. L. 1999. *Entwined Lives: Twins and What They Tell Us about Human Behavior*. Dutton/Penguin Books.

Segal, Nancy L. 1985. "Monozygotic and Dizygotic Twins: A Comparative Analysis of Mental Ability Profiles." *Child Development* 56: 1051–58.

Smith, David M., Walter E. Nance, Ke Won Kang, Joe C. Christian, and C. Conrad Johnston, Jr. 1973. "Genetic Factors in Determining Bone Mass." *The Journal of Clinical Investigation* 52(11): 2800–08.

Starr, B. 2008. "Ask a Geneticist." The Tech, Stanford School of Medicine. http://www. thetech.

org/genetics/ask. php.

Stunkard, A. J., T. T. Foch, and Z. Hrubec. 1986. “A Twin Study of Human Obesity.” *Jama* 256(1): 51–54.

Table 1. Impact of One-Child Policy Fines on the Reported Birthrate of Twins in China, 1965-2001

	(1)	(2)	(3)	(4)	(5)
	Full sample	Parents Han	Either Parent Minority	Full sample	Full sample
Variable	Reported Twinning Birth (Yes = 100)				
Policy fine rate (years of local household income)	0.066*	0.072*	0.011	0.022	
	(0.038)	(0.041)	(0.029)	(0.035)	
Policy fine rate * parents Han				0.052***	
				(0.016)	
Born after 1979 * parents Han					0.154***
					(0.030)
Observations	6,071,870	5,654,203	417,667	6,071,870	6,071,870
R-squared	0.001	0.001	0.001	0.001	0.001

Note: The dependent variable is whether the birth is twins or not (yes = 100). The data are from the Census 1982, 1990, 2000, and 2005. Coefficients should be interpreted in percentage because the dependent variables in all columns have been multiplied by 100. Covariates include dummies for residency type, province, birth order, mother's education level, mother's age at birth, year of birth, survey year, and interactions between year of birth and survey year. The parents' ethnicity dummy is controlled for in columns 1, 4, and 5. Sampling weights are applied. Robust standard errors in parentheses are clustered at the provincial level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2: Heterogeneous Impacts of the Policy Fine on Reported Birth of Twins, by Type of Residence and Birth Order

	(1)	(2)	(3)
	Parents Han	Subsamples by type of residence	
		Urban	Rural
Variable	Reported Twinning Birth (Yes = 100)		
Panel A: Using the policy fine only as the key independent variable			
Policy fine rate (years of local household income)	0.072*	0.060	0.084*
	(0.041)	(0.041)	(0.042)
Panel B: Interacting the policy fine with dummies for birth order to check heterogeneity			
First birth * policy fine rate	0.039	0.056	0.033
	(0.041)	(0.044)	(0.041)
Second birth * policy fine rate	0.151***	0.093	0.168***
	(0.046)	(0.057)	(0.047)
Third or above birth * policy fine rate	0.072	0.022	0.088*
	(0.045)	(0.060)	(0.043)
Observations	5,654,203	1,422,624	4,231,579

Note: The dependent variable is whether the birth is twins or not (yes = 100). The data are from the Census 1982, 1990, 2000, and 2005. Column 1 restricts the sample to those births with parents of Han ethnicity; columns 2 and 3 further divide the sample into urban and rural subsamples. Coefficients should be interpreted in percentage because the dependent variables in all columns have been multiplied by 100. Covariates include dummies for residency type, parents' ethnicity, province, birth order, mother's age at birth, year of birth, survey year, and interactions between year of birth and survey year. Sampling weights are applied. Robust standard errors in parentheses are clustered at the provincial level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Difference-in-Differences Estimation for the Impact of thE One-Child Policy on the Age Gap between First and Second Births

	(1)	(2)	(3)
	Full sample	Parents Han	Either parent minority
Variable	Age gap between first birth and second birth (in years)		
Twinning in second birth *	0.078***	0.077**	0.100
born after 1980	(0.027)	(0.028)	(0.108)
Twinning in second birth	0.188***	0.182***	0.288***
(Yes = 1)	(0.016)	(0.016)	(0.096)
Observations	1,822,396	1,690,608	131,788
R-squared	0.467	0.471	0.460

Note: The dependent variable is the age gap between the first and second births (years). The data are from the Census 1982, 1990, 2000, and 2005. The sample is restricted to second births. Covariates include residency type, province, birth order, mother's education level, mother's age at first birth, year of birth, survey year, and interactions between year of birth and survey year. Sampling weights are applied and robust standard errors in parentheses are clustered at the provincial level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Impact of the One-Child Policy Fine Rate on Height Difference Between Reported Twins

Variable	(1) Height Difference (cm)	(2) Height Difference (cm)	(3) Height Difference/Mean	(4) Height Difference/Mean
Panel A: Full sample				
Policy fine rate (years of local household income)	1.85***		1.71***	
	(0.31)		(0.23)	
Policy fine rate * same-gender		2.10***		1.79***
		(0.38)		(0.26)
Policy fine rate * different-gender		0.69**		1.33***
		(0.25)		(0.27)
Observations	72	72	72	72
R-squared	0.63	0.64	0.63	0.63
Panel B: Only the latest wave is kept				
Policy fine rate (years of local household income)	3.29**		2.71***	
	(1.08)		(0.59)	
Policy fine rate * same-gender		3.37*		2.59**
		(1.56)		(0.83)
Policy fine rate * different-gender		2.92		3.21**
		(1.94)		(1.15)
Observations	33	33	33	33
R-squared	0.74	0.74	0.77	0.77

Note: The dependent variable for columns 1 and 2 is height difference (cm) and for columns 3 and 4, height difference/mean. The data are from the China Health and Nutrition Survey. The sample consists of twins born between 1979 and 2001. Each observation is derived from one pair of twins. Covariates in all regressions include age and age squared of the twins, mother's age at childbirth, and indicators for same-gender twins, urban residence, whether the boy is taller in different-gender pairs, region, year of birth category, and survey year. Standard errors in parentheses are clustered at the provincial level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix Table A1: Summary statistics

Variables	(1) Full sample	(2) Parents are Han	(3) Either parent is minority
Twinning Rate (%)	0.58	0.59	0.43
	(7.58)	(7.65)	(6.56)
Rural area (Yes = 1)	0.73	0.72	0.81
	(0.45)	(0.45)	(0.40)
Both parents Han ethnicity (Yes = 1)	0.93		
	(0.26)		
Age in years	8.04	8.08	7.62
	(4.67)	(4.67)	(4.62)
Mother's age at childbirth	23.25	23.28	22.82
	(2.97)	(2.95)	(3.18)
Birth order			
First (Yes = 1)	0.57	0.57	0.51
	(0.50)	(0.50)	(0.50)
Second (Yes = 1)	0.30	0.29	0.32
	(0.46)	(0.46)	(0.46)
Third or above (Yes = 1)	0.14	0.14	0.18
	(0.35)	(0.34)	(0.38)
Observations	6,071,870	5,654,203	417,667

Note: The data are from the Census 1982, 1990, 2000, and 2005. The sample is restricted to births before 2001. Standard deviations are in parentheses.

Appendix Table A2: Summary statistics in CHNS

Variable	(1) All twins	(2) By type of twins: Same gender	(3) By type of twins: Different gender
Height gap (cm)	2.18	1.30	4.65
	(2.74)	(1.59)	(3.68)
Mean height (cm)	121.15	119.58	125.55
	(27.59)	(28.47)	(25.17)
Gap/mean ratio in percent	1.75	1.02	3.78
	(2.14)	(1.14)	(2.90)
Observations	72	53	19

Note: The data are from the China Health and Nutrition Survey. The sample is restricted to births before 2001. Standard deviations are in parentheses.

Appendix Table B1: One-Child Policy Fine Predicted by the Prior Rate of Twin Births, Post-1980

	(1)	(2)	(3)
	One-Child Policy Fine Rate (Years of local household income)		
Variable	1 year later	3 years later	5 years later
Twinning birth (Yes = 1)	0.013	0.005	-0.002
	(0.008)	(0.005)	(0.006)
Observations	3,846,783	3,617,403	3,412,645
R-squared	0.754	0.778	0.800

Note: The dependent variable the One-Child Policy fine rate in years of local household income. The data are from the Census 1982, 1990, 2000, and 2005. Post-1980 births are used because the One-Child Policy fine started in 1979 and provincial governments started to have local policies in 1984. Micro-level data are used to test whether contemporaneous twin births in local provinces have predictive power on the fine rates in the future, by regressing the fine rates in the next one, three, or five years on the indicator of twinning in the birth-level data, respectively, in columns 1, 2, and 3. The control variables in all columns are the same as those in column 1 of Table 1. They include continuous variables, such as the provincial time trend, and indicator variables, such as residency type, parents' ethnicity, province, birth order, mother's age at first birth, year of birth, survey year, and interactions between year of birth and survey year. Sampling weights are applied. Robust standard errors in parentheses are clustered at the provincial level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

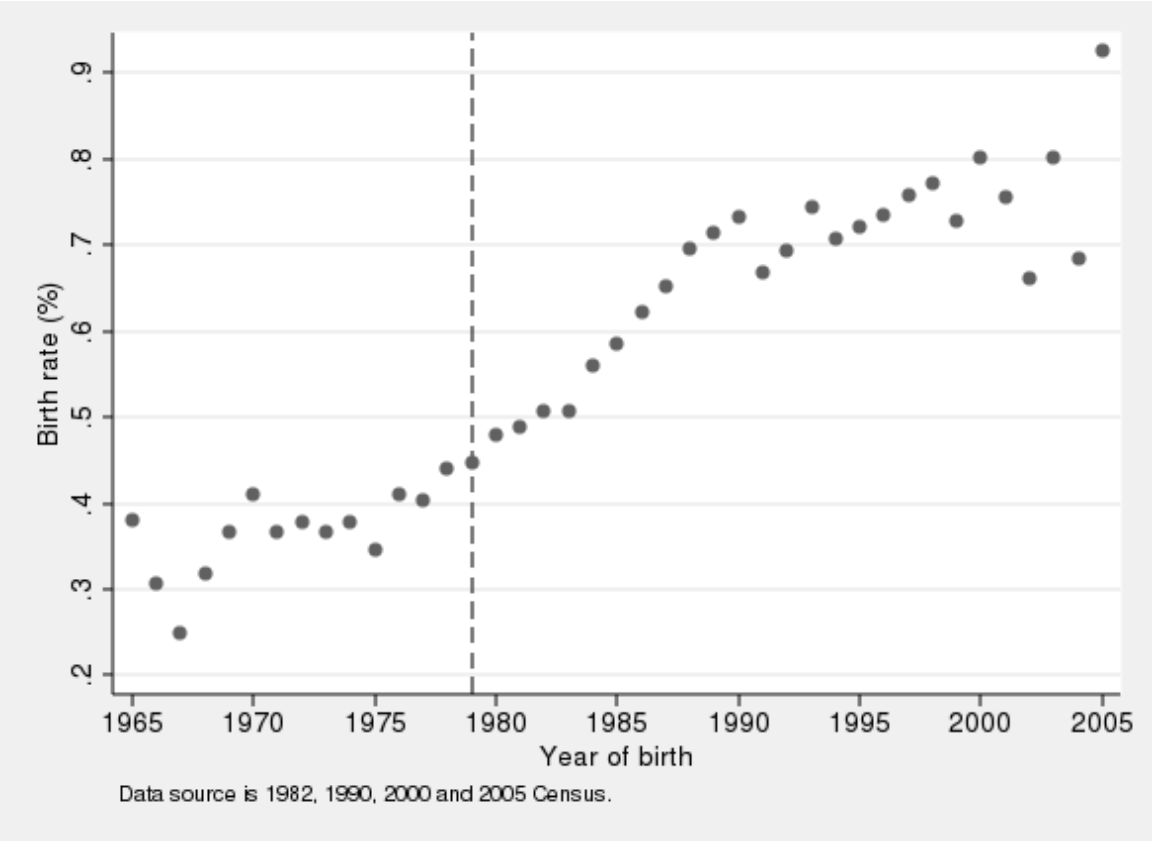


Figure 1. Twining Birth Rate against Year of Birth, 1965–2005

Note: Data are from Census 1982, 1990, 2000, and 2005. Twinning rates in each birth cohort are plotted against year of birth. The vertical dashed line marks 1979, when the One-Child Policy was formally introduced.

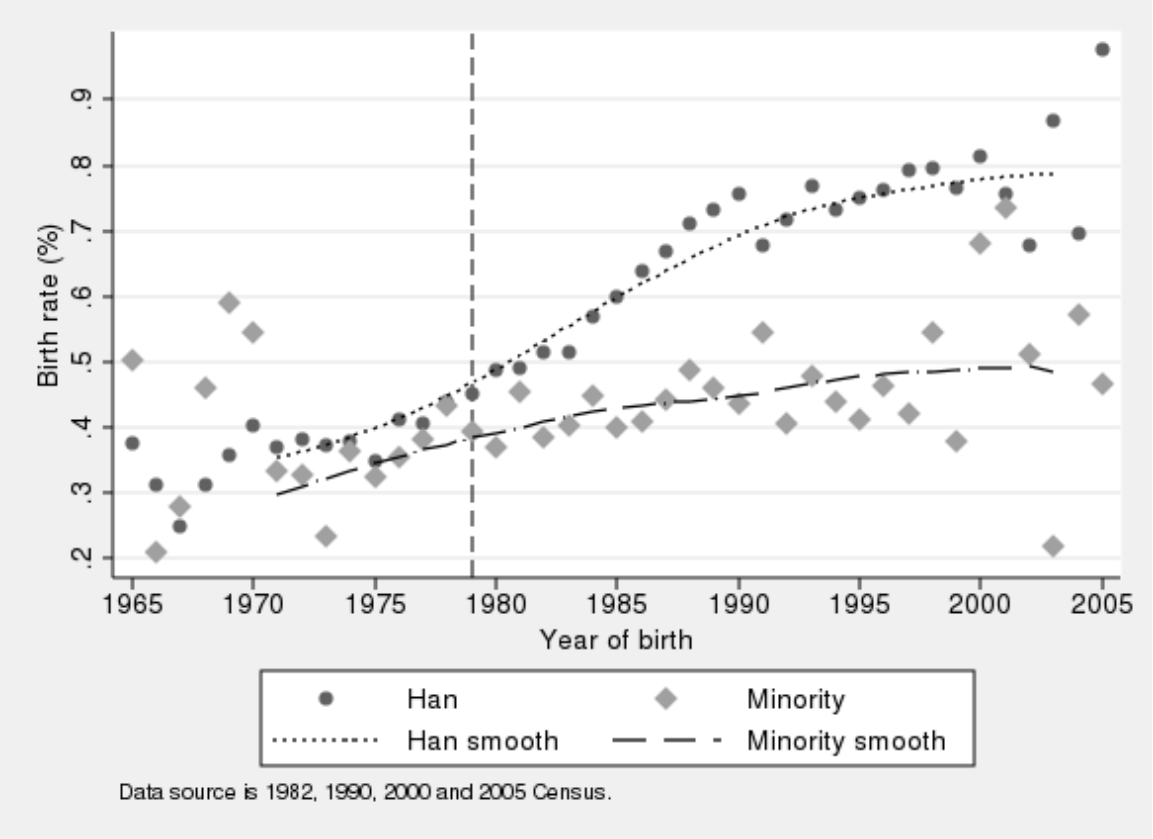
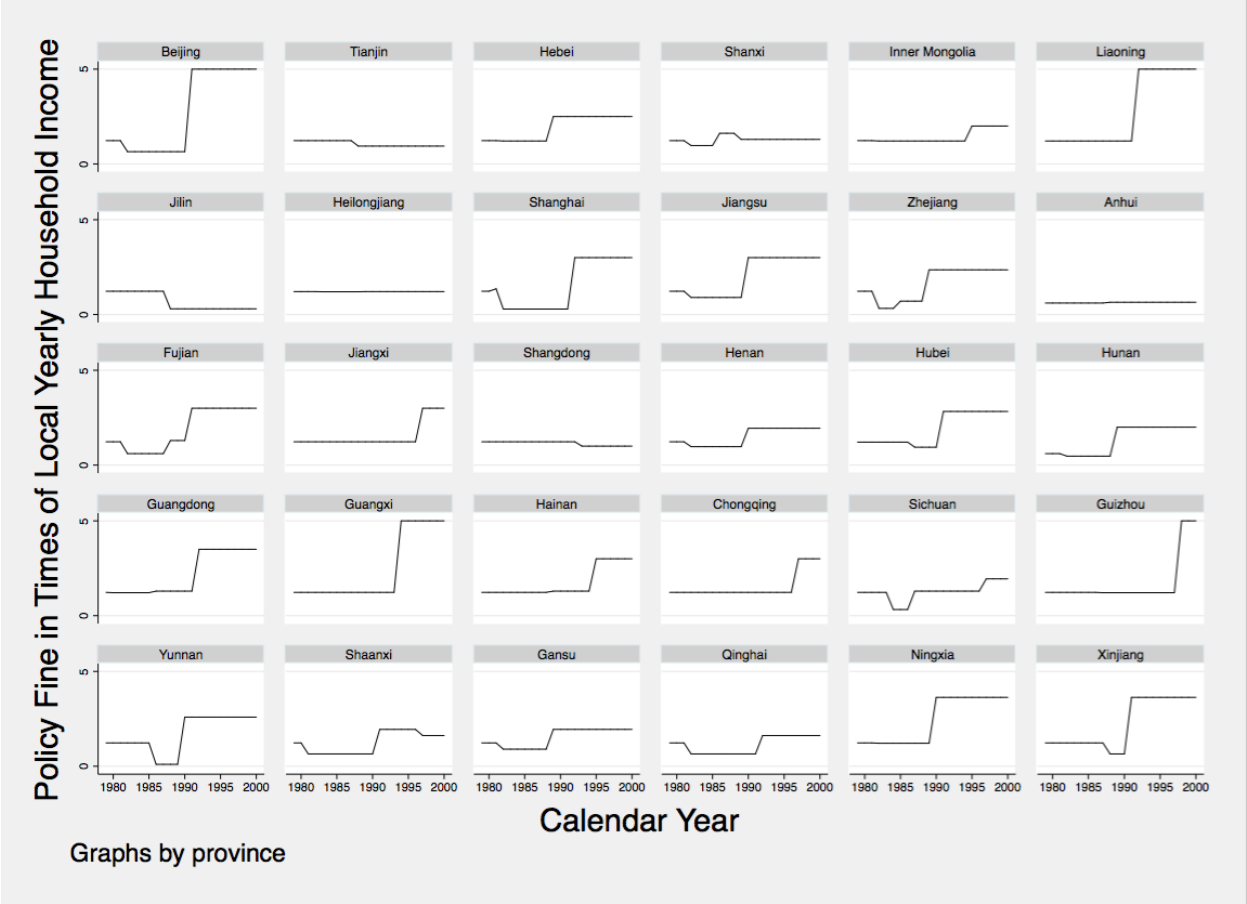


Figure 2. Twinning Rate against Year of Birth, by Ethnicity, 1965–2005

Note: The figure plots the twin birthrate by the ethnicity group of the parents against the year of birth. The dashed lines plot LOWESS-smoothed trends with bandwidth 0.8. The figure shows that the trends in the two groups are almost identical prior to the introduction of the One-Child Policy. However, the difference between Han and minority ethnicity increases after the introduction of the One-Child Policy in 1979 (the vertical dashed line). The increased difference is partly because the One-Child Policy mainly restricts the fertility of families of Han ethnicity rather than minorities, which motivates the Han to "make" twins.



Appendix Figure A1. One-Child Policy Fine Rates, by Province, 1979–2000

Note: This figure plots the average monetary penalty rate for one unauthorized birth from 1979 to 2000 in each province. The data are from Ebenstein (2010). The One-Child Policy regulatory fine (policy fine) is formulated in multiples of annual local household income.