Too Early or Too Late:
What have we Learned from the 30-Year Two-child Policy Experiment

This Version: March 23, 2016

Yu Qin
rstqyu@nus.edu.sg
Department of Real Estate
National University of Singapore

Fei Wang
fwang@ruc.edu.cn
School of Labor and Human Resources
Renmin University of China

Abstract

In January 2016, China ended its 35-year-old one-child policy and replaced it with a nationwide two-child policy. However, it remains unclear whether a two-child policy can effectively increase the fertility level in China. In this paper, we review the 30-year two-child policy experiment that has been carried out in Yicheng, Shanxi, to assess the impact of this policy on the crude birth rate. We adopted a synthetic control approach, which allowed us to conduct a rigorous counterfactual analysis. We failed to find any short-term impacts of the two-child policy on the Yicheng birth rate prior to the 1990s. In the long run, we estimate that the two-child policy will lead to a maximum of 2 more births per 1,000 people every year compared to the one-child policy.

JEL Code: J13, J18

1. Introduction

To tackle the problem of aging, an increasing number of countries have been adopting population policies designed to increase birth rates (United Nations 2013). The effects of such policies are mixed; however, in East Asia they tend to fail more often than in other regions (McDonald 2006). China, the most populous country in the world, has followed a recent trend and began allowing married couples to have two children on Jan 1, 2016, ending a 35-year long one-child policy (OCP). Assessing the potential effects of China's policy change on fertility will not only add to the body of pro-natal policy literature but also has implications for future policy adjustments in China.

It is still far too early to evaluate the new policy directly; however, its effects can be estimated by investigating people's desires to have a second child or by exploring the impact of other two-child policies that have been implemented in China's recent history. Evidence based on the desire for a second child was neither consistent nor credible.[1] Wang, Hu and Zhang (2013) simulated a scenario in which a two-child policy was implemented in 2015, and found that the total fertility rate for 2016 would, at most, increase by 0.5, or approximately 6 million births. However, Zhai, Zhang and Jin (2014) found that the total fertility rate would increase to 3.0 with the institution of a two-child policy, which represents almost 20 million more births in the first year of the new policy.

In the 1970s, before the one-child policy (OCP) was implemented, China allowed all couples to have two children. After the implementation of the nationwide OCP, two children were permitted if a couple was from an ethnic minority, or if they were from a rural region and their first child was a girl. Because the two-child policies are decades old or only applicable to specific populations, we must look to a unique policy experiment implemented in Yicheng, Shanxi Province (location depicted in Figure 1) to see how China's new policy might affect current birth rates. Starting in 1985, Yicheng was granted an exception from the OCP. The locality was designated an experimental zone in which almost all couples were given the option to have two children. This unique experiment provides an excellent opportunity for scholars to investigate the potential consequences of a two-child policy using historical data.

Gu and Wang (2009), Wei and Zhang (2014), and Wu (2014) have all compared the demographics in Yicheng both before and after the experiment and to other counties in the same prefecture city; and they each concluded that replacing the OCP with a two-child policy had little positive impact on fertility levels.

However, it has been challenging to quantitatively estimate the impact of the two-child policy in Yicheng in an unbiased way. On the one hand, before and after comparisons of Yicheng's birth rate may not accurately reflect the true effects of the population policy given possible changes in other determinants of fertility rates. Likewise, it is difficult to carry out a traditional difference-in-differences (DID) analysis, which would account for the differences between before and after birth rates in reference jurisdictions, for two reasons. First, any inference based on a DID analysis is likely to be biased if the number

[1] Desire to have a second child often fails to predict actual behaviors (e.g. Adsera 2006).

of treatment units is small. Second, the control and treatment units analyzed must have parallel growth trends associated with the outcome variables prior to the policy experiment. This is referred to as the "parallel trend assumption" in the DID framework. However, our data indicates that the birth rates in Yicheng and the studied control counties had significantly different growth patterns prior to the experiment, which violates this assumption.

In this paper, we use the synthetic control approach to re-examine Yicheng's two-child policy experiment and more accurately estimate its impact on the locale's fertility rate. The synthetic control method is most suitable for comparative case studies in which only one or a few treated units exist. Positive weights are assigned to a number of control units from a donor pool of counties in the same province such that the weighted average birth rates of the selected control counties best mirror Yicheng's birth rate trend prior to the two-child policy; and the weighted averages of the birth rate determinants from the control counties also match their counterparts in Yicheng prior to the policy's implementation. The formation of synthetic control units into a study group provides for a rigorous counterfactual analysis to more accurately evaluate the policy's effectiveness.

Comparing the crude birth rate in Yicheng to a "synthetic Yicheng", we find that from 1985-1990, the first six years during which the two-child policy was implemented, the birth rate in Yicheng was not significantly different from that in other counties and districts within the same province. This is likely because both the two-child policy in Yicheng and the OCP in other counties were weakly enforced during that time. Nevertheless, the impact of this policy becomes apparent in the long run, once the policies were strictly enforced. The 2000 and 2010 population censuses reveal CBR differences of 2.35 and 1.97 for Yicheng and synthetic Yicheng, respectively, indicating that the two-child policy would lead to a maximum of 2 more births per 1,000 people every year compared to the OCP.

This paper is one of the first to present a rigorous counterfactual analysis of the potential impact of relaxing the OCP on birth rate using a two-child policy experiment, and it is likely to become an important reference for the further relaxation of the OCP nationwide. Our analysis suggests that replacing the OCP with a two-child policy may have little impact on crude birth rate in the short term and would have a rather limited impact, if any, compared to prevailing estimates, in the long term. Given that our estimate is likely to be an upper bound of the true impact, the relaxation of the OCP may have a very limited impact on China's overall birth rate and fertility level.

The rest of the paper is organized as follows: Section 2 provides the background for Yicheng's two-child policy experiment; Section 3 presents a conceptual framework to through which to consider the potential impact of the OCP's relaxation; Section 4 describes the data used in this paper; Section 5 introduces the empirical strategy; Section 6 presents the main findings; Section 7 describes the robustness checks performed; and Section 8 concludes.

2. Policy Background

China began to implement its OCP in 1980. Under this policy, a married couple was generally allowed to have, at most, one child.[2] However, the OCP was difficult to enforce nationwide, especially in rural areas where the policy threatened to significantly reduce the household labor force and decrease a family's capacity to support agricultural production. Families in rural areas were also more likely to express a desire for male children, and birth controls reduced a couple's chances of having a son. Given the realities of implementing the OCP in rural areas, Chinese authorities began to relax the OCP in the mid-1980s, gradually allowing rural married couples to have a second child if the first child was female, so long as the couple waited a sufficient amount of time between the two births (Yang 2004, pp. 136-137).[3]

Some scholars proposed alternative solutions to this caveat. In the spring of 1984, Liang Zhongtang, a demographer in the Shanxi Province People's Government Economic Research Center at that time, proposed a restricted two-child policy that would require couples to marry later and wait longer before having a second child. In 1985, Liang's advocacy for this policy convinced the provincial government to allow Yicheng, a county in Linfen prefecture city, to replace the OCP with the two-child policy proposed by the demographer. Figure 1 shows the location of Shanxi province in China and the location of Yicheng in Shanxi. Shanxi is in central North China, and Yicheng is located in South Shanxi.

In fact, Yicheng was not chosen for this experiment at random. In an interview,[4] Liang summarized the three reasons for why Yicheng was selected as a pilot for this experiment. First, given that it was more difficult to enforce the OCP in rural areas compared to cities, it was logical to test the two-child policy in a county with a primarily rural population. In 1985, over 90% of Yicheng's population lived in rural areas. Second, Yicheng had railroad access; therefore, program implementation and monitoring would be more easily executed. Lastly, the experiment was welcomed and supported throughout Yicheng. In our research, we try our best to control for the demographic and socioeconomic variables that are correlated with the placement of the program.

Yicheng's two-child policy includes the following measures: 1) all couples are encouraged to delay marriage, to postpone parenthood and to have fewer children; 2) the "one child per couple" norm is enthusiastically promoted; 3) rural one child families are offered financial incentives (i.e., financial rewards and preferential access to education and health services); and 4) state employees and urban couples are limited to one child only, except under special circumstances. Rural couples fulfilling the following requirements are allowed to have two children: a) they must marry three years later than the minimum marriageable age, as specified in The Law of Marriage (men at 22 and

[2] For additional details about the OCP and its earlier policy versions, as well as their effects on birth rate, see Wang (2014).

[3] Urban and rural residents whose first child was a male were generally subject to the one-child birth quota as before. Therefore, we still refer to the policy as the OCP.

[4] Jiao, H. (2012, August 21). Interview Liang Zhongtang, a Family Planning Expert and Demographer. *Legal Weekly*. Retrieved September 19, 2015, from http://www.legalweekly.cn/index.php/Index/article/id/718.

women at 20); (b) the wife bears her first child at 24 and the second at 30[5] (the birth spacing requirement was adjusted from 6 years to 4 years in 2007); (c) the wife must apply for a birth permit for her second child and wait for a quota; and (d) the couple must use effective contraception after their first child is born and must accept sterilization after their second child. Births beyond two children are strictly prohibited without exception and are subject to financial and disciplinary sanctions (Wei and Zhang, 2014).

There are limited references related to the enforcement of Yicheng's two-child policy experiment. In the first decade of Yicheng's experiment, more than half of the second children born were birthed without a permit because the parents violated the late marriage requirement or had not waited long enough between children to conceive. However, the ratio of unpermitted second children significantly dropped after 1995 and reached almost zero in 2010 as late marriages and larger gaps between pregnancies became more prevalent (Wu 2014).

Yicheng was the first county to implement the two-child policy experiment, but it was not the only one. A dozen more pilot counties, including one jurisdiction in Shanxi (i.e., Xinrong district in Datong prefecture) and a number of counties in other provinces also adopted this law. However, starting in the late 1980s, the two-child policies in most of the pilot regions were replaced by the hybrid one-child policy in which a rural couple could have two children only if the first child was female. Gu and Wang (2009) noted that the two-child policy was maintained in four regions: Chengde in Hebei Province, Enshi in Hubei Province, Jiuquan in Gansu Province, and Yicheng in Shanxi Province. Nevertheless, Liang (2014b) has argued that Yicheng is the only appropriate location in which to study the two-child policy because Chengde's policy was only enforced in the least developed mountain areas; Enshi has a large ethnic minority population, so the policy effects there may not be representative; and Jiuquan's policy was often interrupted and remained illegal during most of the 1990s. Therefore, our research focuses on the two-child policy in Yicheng.

3. Conceptual Framework

This paper estimates the impact of a two-child policy on population growth compared to an OCP at the aggregate (county) level. The only measure of county fertility levels available to us is crude birth rate (CBR, i.e., the number of births per 1,000 people in a year). In contrast to total fertility rate as adjusted for age of the mother, the CBR of a county is likely to be affected by the gender and age structure of the county (Easterlin 1978). For example, more females of child-bearing age in a county could lead to higher CBR. Therefore, we need to collect information on the gender and age distributions of selected county for our analysis.

In addition to gender and age distribution, Easterlin and Crimmins (1985) discuss three factors that may determine fertility levels: the demand for children, the supply of children and the costs of fertility regulation.

[5] Women who only want one child do not have to marry or give birth late.

The demand for children is largely driven by socioeconomic factors. At the county level, we consider the proportion of the rural to urban population, economic conditions and education in our analysis. The supply of children is mainly determined by people's fecundity and the infant mortality rate. Fecundity cannot be directly measured at the county level but can be broken down into measurable determinants, such as the age structure, economic conditions and health variables. These variables are related to nutritional and biological factors that govern fecundity (Frisch 1982; Wood 1989). Infant mortality rate data are unavailable but can be approximated by health factors.

The costs of fertility regulation include people's attitudes about family planning, the accessibility of fertility control methods and supplies, and penalties for violating family planning policies. We assume that the differences in the costs of fertility regulation between Yicheng and other counties are primarily accounted for in their different family planning policies. The policies also differ between urban and rural areas, and between ethnic majority Han communities and minority communities (Wang 2014). Therefore, the geographic and ethnic makeup of the population also needs to be considered so that the differences in policy intensities between Yicheng and the other counties studied are solely related to those counties' distinct policies rather than the geographic and ethnic distribution of the population. Thus, we must consider the share of the population that lives in rural communities and the ethnic distribution of the population by county in our analysis. Unfortunately, county level data on the ethnic makeup of the population are not available; nevertheless, this is unlikely to be a problem in our study as Shanxi province has been home to very few ethnic minorities in the past decades.[6]

In addition, people's tastes and culture may also affect birth rates, both in China (Arnold and Liu 1986) and in other countries (Fernández and Fogli 2006). We consider these factors in our analysis either by using the best available observed variables as proxies or by assuming that unobservable characteristics are similar between Yicheng and other counties, where appropriate.

4. **Data Sources**

As stated above, the primary variables of interest are county level CBR and its determinants in Shanxi province. We restrict our data to the counties in Shanxi for two reasons. First, the number of counties in Shanxi is large enough to perform a comparative case study; and second, unobservable characteristics are likely to be more similar at the county level for counties within the same province.

We obtained CBRs and crude death rates (CDR, number of deaths per 1,000 people in a year) to calculate a health variable for all 116 counties and districts in Shanxi from 1949 to 1990 (from the published book "40 years of population in Shanxi: 1949-1990"). The data availability varied across jurisdictions. The shortest panel contains population measures from 1972 to 1990, while the longest records both variables continuously from

[6] The 1953 Chinese census showed that the percentage of ethnic minorities in Shanxi was only 0.14%. By the 2010 census, the figure had risen slightly to 0.26%. The percentages have been similarly low in almost all counties within Shanxi.

1949 to 1990. We excluded Xinrong district in Datong prefecture from our sample because Xinrong was one of the localities that implemented a two-child policy during the late 1980s and early 1990s, which may have an impact on the results. In addition, we adjusted our data accordingly to correct the number of administrative units analyzed during our sample period.[7] To assess the long term impact of Yicheng's two child policy, we also collected the CBR and CDR for all counties in Shanxi from the 2000 and 2010 population censuses.[8]

In addition to CDR, we required other predictors of CBR. Based on the literature discussed in the above section, we included the following sets of variables: 1) population by age cohort (age 0-14, 15-59 and 60 and above), and gender for each prefecture from the 1982 population census. As the age by gender data were not available at the county level, we assumed that the age-gender distribution in each county was the same as the prefecture it was affiliated with.[9] In addition, we used population by gender data for each county from 1949-1990 to complement the prefecture level age-gender data, as per "40 years of population in Shanxi: 1949-1990". These variables were used to control the variations in CBR that occur as a result of the differences in the age-gender distribution; 2) other indicators include the rural proportion of the population, GDP per capita, rural personal income, number of elementary schools per 1,000 people, number of middle schools per 1,000 people and number of hospitals per 1,000 people. For districts and counties in Linfen prefecture, with which Yicheng is affiliated, we collected data on these variables at the county level from "50 years of Linfen", a publication that contains yearly statistics for all jurisdictions in Linfen prefecture. For other counties and districts, we were only able to obtain the rural proportion of the population at the county level from "40 years of population in Shanxi: 1949-1990".[10,11] The rest of the variables were only available at the prefecture level and were obtained from "60 years of Brilliant Shanxi province".[12] For these variables, we assumed that they had the same value at the county level as the prefecture the county was affiliated with.

5. **Empirical Strategy**

[7] We removed Yanbei District from our sample as it was disbanded in 1993, and we reassigned the ten affected counties to their corresponding prefectures based on the current administrative hierarchy. Specifically, we assigned Tianzhen, Yanggao, Guangling, Lingqiu, Hunyuan, Zuoyun and Datong counties to Datong prefecture; and Huairen, Ying and Youyu counties to Shuozhou prefecture.

[8] There were a large number of administrative changes between the 2000 and 2010 population censuses compared to earlier data, especially in the urban districts. If the administrative boundary did not change, we simply matched the new jurisdiction name with the old name. In addition, we made the three further changes: 1) we matched the weighted average of Jinyuan district and Xiaodian district in Taiyuan prefecture to the old Nanjiao district in Taiyuan; 2) we matched the weighted average of urban and suburban districts in Changzhi prefecture to the old urban district in Changzhi; and 3) we matched the urban district and Zezhou county in Jincheng prefecture to the old urban district in Jincheng.

[9] The ten counties administered by Yanbei district, which was later disbanded, were assigned values based on Yanbei district.

[10] The data for the urban districts in Taiyuan, Yangquan and Datong are missing. We inferred these values by using the rural proportion of the population for the entire prefecture as a proxy.

[11] The data for Gujiao city in Taiyuan are missing. We inferred these values based on the rural proportion of the population in Taiyuan prefecture and the other jurisdictions in Taiyuan.

[12] The data in the three books, "40 years of population in Shanxi: 1949-1990", "60 years of Brilliant Shanxi province", and "50 years of Linfen" are all sourced from local statistical and public security authorities.

Figure 2 illustrates the CBR trends in Yicheng and 114 other districts and counties in Shanxi province. These data clearly suggest that the CBR in Yicheng varies greatly from that in the rest of the province. First, the CBR in Yicheng was significantly below the average CBR in other jurisdictions. In addition, the CBR growth differed in Yicheng compared to other jurisdictions. A formal test of the differences between Yicheng and other jurisdictions was carried out using a DID analysis. Using the data for 115 counties from 1972-1990, 2000, and 2010, we performed a regression analysis of CBR in relation to county dummies, year dummies, and the interactions of Yicheng and year dummies.[13] Figure 3 reports the 95% confidence band for the coefficients of the interactions, with year 1984 as the baseline year. Almost all of the coefficients were significantly negative before year 1985, indicating a significantly larger variation in CBR between Yicheng and other counties in these years compared to 1984. In addition, any pair of coefficients in adjacent years prior to 1985 were significantly different from one another, which clearly suggests different growth patterns for Yicheng and the rest of the province. Therefore, the estimated impact of the pilot two-child policy is likely to be biased if we directly implement a DID analysis. Moreover, Cameron and Miller (2015) indicate that a DID model is likely to be inconsistent if there are too few treatment groups compared to control groups, which is the case in our study.

As difference-in-differences cannot appropriately identify the policy impact of Yicheng's experiment, we adopted a synthetic control approach, which sufficiently addresses the shortcomings of a traditional DID analysis. The synthetic control method (see Abadie and Gardeazabal, 2003; Abadie et al., 2010 and Abadie et al., 2015) allows us to construct an artificial control group that almost exactly mimics the growth patterns of the treated unit prior to the policy experiment.

Assume that there are *J* units. The first unit, *J*, is the treated unit, and *J-1* units are used to construct a synthetic control unit that is comparable to the treated unit. Assume Y_{it} is the dependent variable for unit *i* over period *t*. Y_{it} is determined by the following factor model for $i = 1,2 \ldots J$ in the pre-treatment periods:

$$Y_{it} = \delta_t + \boldsymbol{\theta}_t \boldsymbol{Z}_i + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \varepsilon_{it}.$$

This factor model allows time-varying effects for both observable $\boldsymbol{Z}_i$, and unobservable characteristic, $\boldsymbol{\mu}_i$, which is less restrictive than traditional DID models with fixed effects. Suppose there is a set of nonnegative weights $(w_2, w_3, \ldots, w_J)$ and $\sum_{j=2}^{J} w_j = 1$, such that

$$\sum_{j=2}^{J} w_j Y_{jt} = Y_{1t}, \text{ for each pre-treatment period}$$

$$\sum_{j=2}^{J} w_j \boldsymbol{Z}_j = \boldsymbol{Z}_1.$$

Given this, Abadie et al. (2010) show that the treatment effect in any post-treatment period *t* would be $Y_{1t} - \sum_{j=2}^{J} w_j Y_{jt}$, as long as the characteristics of the donor pool of counties are not substantially different from those of Yicheng and the pre-treatment period used to match Yicheng to its synthetic counterpart is sufficiently long.[14]

[13] The standard errors are clustered at the county level.

[14] Abadie et al. (2010) also show that, with such conditions, both observables and unobservables are matched between Yicheng and synthetic Yicheng.

In our study, Yicheng is the treated unit, and the other 114 counties in Shanxi form the donor pool that is used to construct the synthetic Yicheng. Weights are assigned to the donor pool units, such that 1) the CBR of Yicheng is as close as possible to the weighted average of the CBR for the donor pool units in each period prior to 1985; 2) for each determinant of CBR, its average value over the pre-intervention periods in Yicheng is as close as its counterpart in the weighted donor pool as possible. With such a set of weights, a synthetic Yicheng is constructed from the donor pool and mirrors Yicheng before the policy intervention in terms of CBR and its determinants.

Our core dependent variable is CBR at the county level. As introduced in previous sections, we included the following variables as predictors. First, we include variables related to gender by age cohort (age 0-14, 15-59, and age 60 and above) population structure. Specifically, we include the share of males and females for the three age cohorts, respectively. It is also worth noting that the sex ratio under age 15 may reflect gender preferences in a given area to a certain extent. Second, we include health measures, such as CDR and the number of hospitals per 1,000 people. Third, we include a series of socioeconomic indicators, including a) the rural proportion of the population, as CBR varies from urban to rural areas; b) GDP per capita and rural income in natural log form, which measures the income level of a region; and c) measures of the average educational level in each region, which is likely to affect CBR and includes the number of elementary schools and middle schools per 1,000 people. Lastly, Abadie et al. (2010) suggest that controlling for lagged values of the dependent variable helps to better fit the dependent variable's trend. Thus, we also include historical birth rates from 1972 to 1984 as predictors.

We use 1972-1984 as the sample period to fit the CBR growth trend between Yicheng and the "synthetic Yicheng", as it is the longest period covered by all 115 counties and districts in Shanxi province.[15] We did not include any jurisdictions outside Shanxi province as family planning policies vary from province to province, making the comparison less relevant. We then use the data from 1985 to 1990 to estimate the short-term impact, and the data in 2000 and 2010 to infer the long-term impact.

6. **Results**

As shown in Table 1, the synthetic control method generates positive weights for Kuangqu in Datong prefecture (0.302), Hexi district in Taiyuan prefecture (0.283), Fushan county in Linfen prefecture (0.195), Qinyuan county in Changzhi prefecture (0.184) and Pianguan county in Xinzhou prefecture (0.036). The "synthetic Yicheng" is similar to Yicheng in many respects, as suggested in Table 2. Particularly, birth rates prior to 1985 appear similar. The simple average birth rates of the entire donor pool are substantially larger than those in Yicheng, but the synthetic Yicheng values match Yicheng values well. However, three predictors are not balanced across the treated and synthetic control group: mortality rate per 1,000 people, number of hospitals per 1,000 people, and the rural proportion of the population. The "synthetic Yicheng" has a lower

[15] Except for Xinrong in Datong which was affected by two-child policy before early 1990s.

mortality rate, more hospitals and a lower rural population. All of these imbalances suggest that the synthetic control group is more urbanized, healthier, and thus likely to have a lower birth rate. Therefore, our method may over-estimate the impact of a two-child policy on birth rate.

Figure 4 presents CBR in Yicheng and its synthetic counterpart over time. In general, the synthetic control group closely follows the trend of birth rates in Yicheng prior to the implementation of the two-child policy in 1985. Both birth rate series show downward trends before 1985, which reflects the effects of both family planning policies and socioeconomic development. The birth rate increased in both Yicheng and the "synthetic Yicheng" from 1985-1990, and then fell in both 2000 and 2010. However, the synthetic control group's CBR appeared to fall at a faster rate than did the CBR of Yicheng.

One possible reason for the inverse-U patterns of both series may be the change in policy enforcement. In Yicheng, as mentioned in previous sections, a large portion of second births that occurred before the mid-1990s were unpermitted, which may have resulted in a fertility boom in the first decade of the two-child policy. After the mid-1990s, unpermitted second births became much less common due to stricter enforcement, which likely explains the downward trend in the county's fertility rate in the 2000s. For other counties, the Chinese central government relaxed the OCP to a one-and-half-child policy in rural areas around the mid-1980s. However, this policy relaxation was not clearly understood, so a number of local authorities and residents believed that two children were generally allowed (Yang 2004, pp. 137). This may explain the rise in fertility rates in the synthetic Yicheng, as well as the similarity in birth rates between Yicheng and its synthetic counterpart in the late-1980s. Shanxi legislated family planning in 1989, after which policy enforcement grew stricter (Yang 2004, pp. 138, 161). This may have contributed to the fertility decline in the 2000s. In other words, the birth rate differences in 2000 and 2010 may better reflect the differences between the actual OCP and the pilot two-child policy.[16] The 2000 and 2010 differences in CBR between Yicheng and synthetic Yicheng are 2.35 and 1.97, respectively, indicating that the two-child policy would result in approximately 2 more births per 1,000 people every year compared to the OCP. Therefore, the long-term effects of this policy change are fairly small.

However, the estimated magnitude of this effect may be somewhat conservative in the *short-run* for two reasons. On the one hand, we cannot precisely capture changes related to the sudden implementation of fertility control in Yicheng immediately after the two-child policy went into effect because the loose enforcement of the OCP in the synthetic Yicheng unit in the late-1980s confounded the data. On the other hand, the suppression of fertility levels in Yicheng in 1985 was likely to be less effective than today because two children were allowed prior to 1980.

Given these two concerns, we re-estimate the potential impact of the two-child policy for the first few years using the following two assumptions: first, we assume that the other counties in Shanxi maintained CBR levels as low as those in 1984; and second, we

[16] Another reason for the rise in fertility rates in the late-1980s has to do with the baby boomers reaching childbearing age in the late-1980s (Liang, 2014a).

assume that the CBR in Yicheng reached its peak, not as recorded in 1990, but immediately following the implementation of the two-child policy. These two assumptions allow us to address the first concern raised above by re-estimating the upper bound of the CBR variation between Yicheng and its synthetic counterpart in the short run to 4.6 births per 1,000 people. Again, this is a rather limited number. Unfortunately, we do not yet have an effective method to address the second concern. However, as fertility declines with age, the impact of suppressed fertility desire on CBR is also likely to decline with age. That said, the fertility peak in Yicheng immediately after 1985 may not differ greatly from what may occur today after the nationwide two-child policy is implemented.

As mentioned in Abadie et al. (2010), our results could be driven entirely by chance; therefore, we ran placebo tests to determine the significance of our estimates. In Figure 5(a), the dark line represents the variation in CBR between Yicheng and the synthetic Yicheng over years. We assumed each of the control units was the treated county, determined the variation in CBR between the county and its synthetic counterpart, and plotted the variation with a light line in Figure 5(a). As a number of control units are badly matched with their synthetic counterparts before the treatment, Figure 5(b) only illustrates those counties whose pre-treatment MSPEs (mean squared prediction error, the average of the squared discrepancies between CBR in a county and in its synthetic counterpart during the period 1972–1984) are not greater than that of Yicheng, reducing the number of counties from 115 to 62. If the variations in Yicheng's CBR and the synthetic Yicheng's CBR are mainly driven by the two-child policy rather than by chance, the dark line should remain above most, if not all, of the light lines.

In both Figure 5 graphs, the dark Yicheng line does not stand out above all of the light lines. However, in 2000 and 2010, the Yicheng values are clearly greater than most of those for the other counties. Table 3 further shows the percentiles for each year's CBR variation after the treatment. According to Table 3, Yicheng's policy effects are not significant between 1985 and 1990. However, in 2000 and 2010, the effects are significant at the 5% and 10% levels.

If the welcoming policy in Yicheng attracted more migrants interested in having more children, the estimated effects may be biased. However, this is unlikely to be a concern for two reasons. First, according to the 1990 census sample, the fraction of migrants in Yicheng did not increase from 1985 to 1990, the first five years of the two-child policy; and it remained low, below 2% in later years. Second, even though the estimated effects may be biased, they tend to be biased toward the high side. In other words, the true effects may be even smaller than estimated, which would not alter the conclusion that the two-child policy will have a limited impact on fertility rates.

7. **Robustness Checks**

Other than the aggregate data collected by local authorities, census samples may also be used to assess the effects of a two-child policy on fertility rates. We confirmed whether the results would be similar if different sources of data were used. Based on a 1% sample of the 1990 Chinese census, we approximated the CBR in Yicheng and other Shanxi

counties from 1972 to 1989. Using a 0.1% sample from the 2000 Chinese census, we approximated the CBR in Yicheng and other Shanxi counties from 1990 to 2000. We also obtained the CBRs of all counties in Shanxi for 2010. Using the synthetic control approach, we assigned weights to all of the other counties in Shanxi such that Yicheng's birth rates from 1972-1984 were as near the weighted birth rates of the other counties as possible. Thus, the variation in birth rate between Yicheng and synthetic Yicheng after 1985 could best be interpreted as the effect of the two-child policy on fertility. Figure 6 shows this policy effect. Given that the county-year sample size from the 2000 census was small, the birth rates were grouped into five-year cohorts from 1990-2000. The figure shows nearly zero short and long term fertility effects from the two-child policy. In other words, we failed to find evidence of strong policy effects on fertility levels using the census samples, which is consistent with the results of the aggregate data obtained from local authorities. Analyses based on census samples raise additional concerns. First, the county-year sample size is small. Therefore, sampling errors would be large and any inferences made rather inaccurate. Second, the calculation of birth rates from census samples assumes a zero mortality rate over time, which may be substantially different from reality. Consequently, estimates made using census samples are merely used to perform robustness checks and should not be considered the primary findings.

The synthetic control approach requires similarities between Yicheng and its donor pool of control counties. Constructing the synthetic Yicheng by only using counties that closely surround Yicheng may have made the assumption more robust. Unfortunately, neither the counties in Linfen prefecture, nor the counties in the neighboring four prefectures of southern Shanxi, formed a well-matched synthetic Yicheng. By including the counties in the four prefectures of central Shanxi, we were able to develop a synthetic Yicheng that fit well with the real-life case study, and the policy effects were similar to those from all of the counties in Shanxi province (Figure 7(a)). Moreover, the robustness check allowed us to confirm that if the county with the largest weight, Kuangqu, in Datong prefecture, is excluded, the results remain stable.

The effectiveness of the synthetic control method must also be evaluated over a sufficiently long pre-treatment period.[17] The results presented in the previous sections are based on matching specific characteristics in Yicheng with those from a donor pool of 114 counties for the period 1972-1984. From 1970-1984, data were available for Yicheng and 108 other counties. Figure 7(b) shows the results from the comparison between Yicheng and the donor pool of 108 counties for the period 1970-1984. The extent of the similarities declines, but is still acceptable.[18] It turns out that the results remain robust if the pre-treatment period is extended by two years. Unfortunately, further extension of the pre-treatment period does not yield well-matched results. Compared to literature, the pre-treatment period used in our study was sufficiently long.[19]

[17] Abadie et al. (2010) indicate that the number of pre-treatment periods should be large enough compared to the scale of the error term in the factor model. As the scale of the error term is immeasurable, there is no clear criterion for determining a proper length for the pre-treatment period.
[18] The MSPE of the matching from 1970-1984 is about 4 times the MSPE of the matching from 1972-1984.
[19] In the classic example of Abadie et al. (2010), the pre-treatment period is 19 years, similar to our best analysis, 15 years.

8. **Conclusion**

In this paper, we study the two-child policy experiment implemented in Yicheng beginning in 1985. We adopt a synthetic control approach, which allows us to conduct a rigorous counterfactual analysis by constructing a "synthetic Yicheng," which mirrored the birth rate in Yicheng prior to the policy experiment. We fail to find any impact of the two-child policy in Yicheng before the 1990s, which may be due to the weak enforcement of the policy in Yicheng and the poor implementation of the OCP in the control counties throughout Shanxi during that time. In the long run, we found that the two-child policy would lead to approximately 2 more births per 1,000 people every year compared to the OCP in Yicheng, which is likely to be an upper bound due to the differences in socioeconomic factors between Yicheng and its synthetic counterpart.

The estimated magnitude of the two-child policy's impact on the birth rate in our paper is lower than estimates predicted by influential demographers, such as Zhai, Zhang and Jin (2014). They conclude that the relaxation of the OCP is likely to lead to an additional 97 million newborns in four years, equivalent to an approximate increase of 18 births per 1,000 people every year. Such population growth would slow the pace of aging in China. In contrast, our analysis suggests that replacing the OCP with a two-child policy will have a very limited impact on population growth, both in the short term and the long term, in rural areas, such as Yicheng. Thus, it will have very limited effects on the age distribution in these areas.

A caveat of this paper is that the estimated impact of the two-child policy presented here cannot be generalized across the country. Instead, the conclusions may be generalized to rural areas similar to Yicheng. At the very least, the 30-year experiment in Yicheng suggests that it is not too early, but may be too late, to alter the aging trend using a two-child policy in such areas.[20]

[20] Simulations by Wang, Zhao and Zhao (2016) have shown that the aging trend cannot be altered even if the birth quota is completely removed.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program." *Journal of the American Statistical Association*105.490 (2010).

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. "Comparative politics and the synthetic control method." *American Journal of Political Science* 59.2 (2015): 495-510.

Abadie, Alberto, and Javier Gardeazabal. "The economic costs of conflict: A case study of the Basque Country." *American economic review* (2003): 113-132.

Adsera, A. (2006). An economic analysis of the gap between desired and actual fertility: The case of Spain. *Review of Economics of the Household*, *4*(1), 75-95.

Arnold, F., & Zhaoxiang, L. (1986). Sex preference, fertility, and family planning in China. *Population and Development Review*, 221-246.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317-372.

Easterlin, R. A. (1978). What will 1984 be like? Socioeconomic implications of recent twists in age structure. *Demography*, *15*(4), 397-432.

Easterlin, R. A., & Crimmins, E. M. (1985). The fertility revolution: A supply-demand analysis. University of Chicago Press.

Fernández, R., & Fogli, A. (2006). Fertility: The role of culture and family experience. *Journal of the European Economic Association*, *4*(2- 3), 552-561.

Frisch, Rose E. "Fertility and population regulation: Demographic implications of the biological determinants of female fecundity." *Social Biology* 29.1-2 (1982): 187-192.

Gu, B., and F. Wang, 2009, An Experiment of Eight Million People (in Chinese), Social Sciences Academic Press (China).

Liang Z., 2014, Studies on China's Fertility Policies (in Chinese), Shanxi People's Press (China).

----------, 2014, China Population Problem (in Chinese), China Development Press (China).

McDonald, P. (2006). Low fertility and the state: The efficacy of policy. *Population and development review*, *32*(3), 485-510.

United Nations. (2013). World Population Policies 2013.

Wang, F.. (2014). *Essays on Family Planning Policies, Doctoral dissertation*, University of Southern California.

Wang, F., L. Zhao and Z. Zhao, 2016, China's Family Planning Policies and Their Labor Market Consequences, IZA discussion paper No. 9746.

Wang, G., Y. Hu and L. Zhang, 2013, Adjustment of Family Planning Policy in China (in Chinese), Social Sciences Academic Press (China).

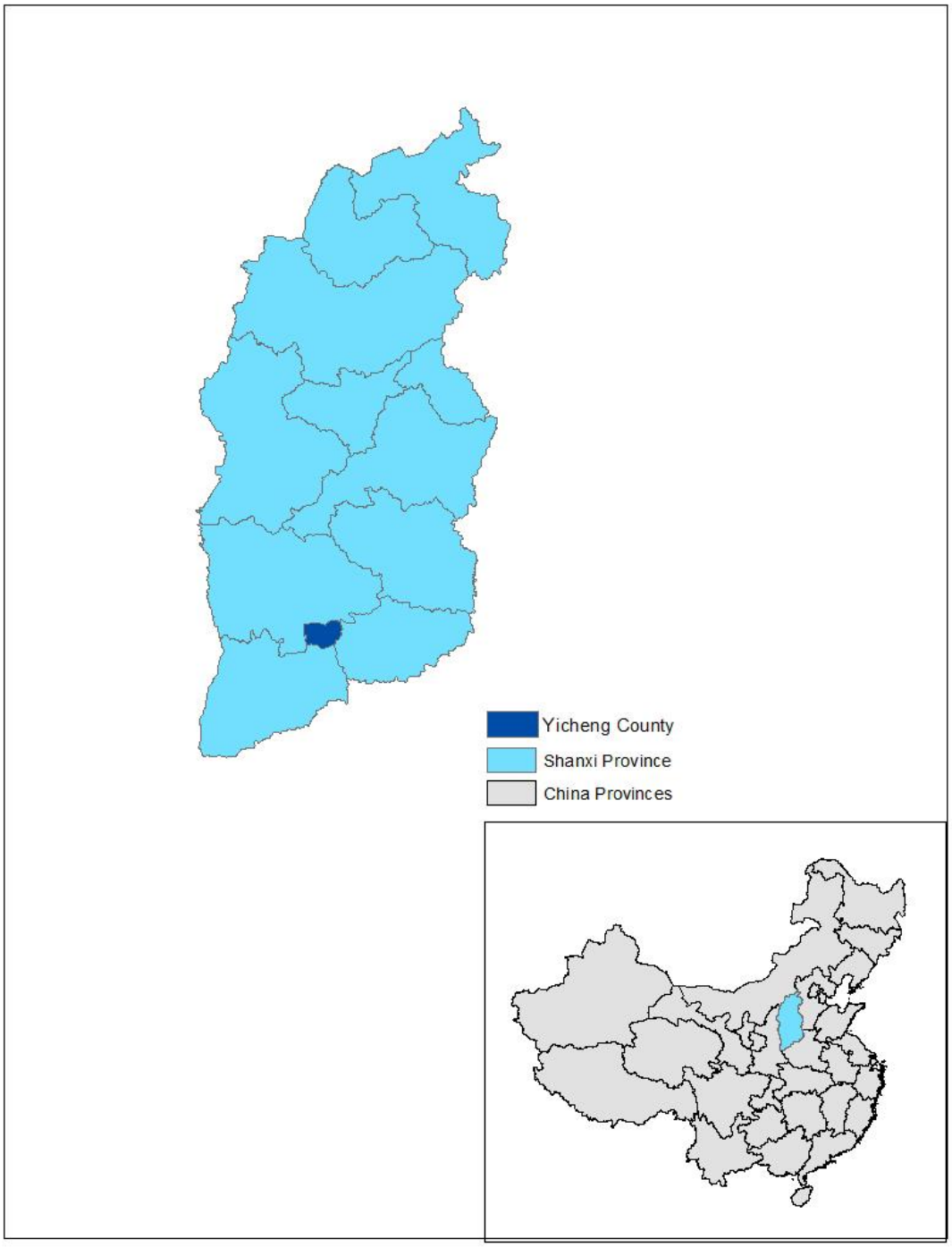
Wei, Yan, and Li Zhang. "Re-examination of the Yicheng Two-Child Program."*China Journal* 72 (2014): 98-120.

Wood, J. W. (1989). Fecundity and natural fertility in humans. *Oxford reviews of reproductive biology*, *11*, 61-109.

Wu, Yanwen. (2014). Preliminary Analysis of A Second-child Experiment on "Later Marry, Later Birth and Longer Interval" in Yicheng County. *Population Journal*, (4), 103-112.

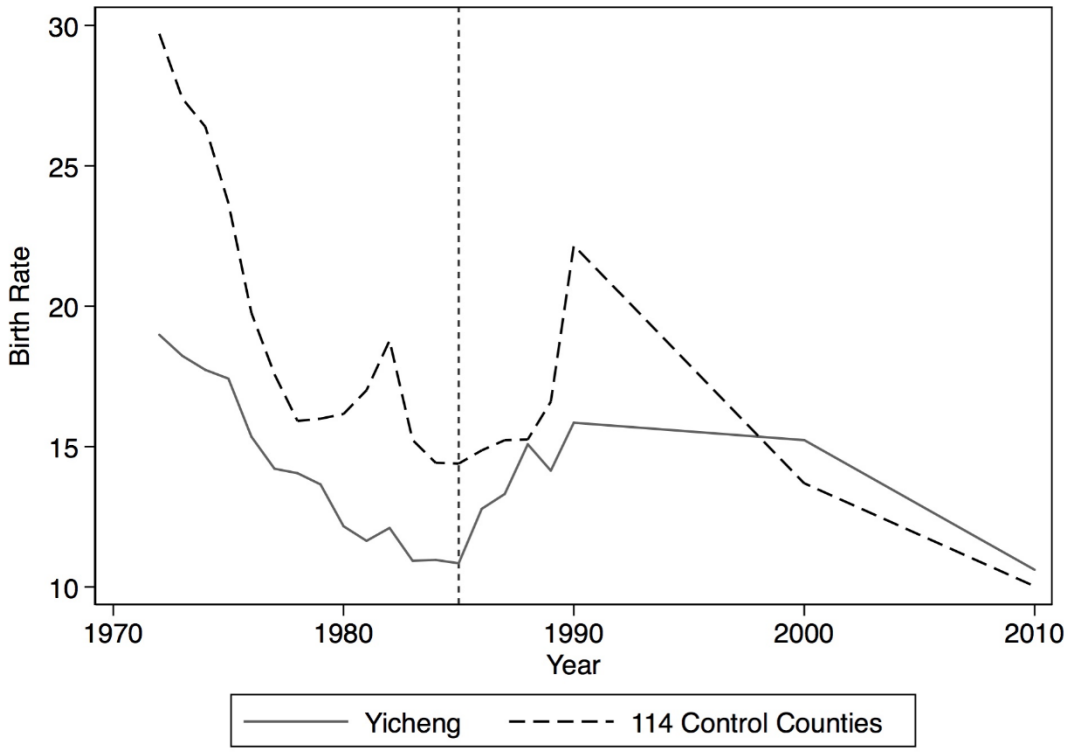
Yang, F. (2004). *Historical Research on Family Planning of Contemporary China,* Doctoral dissertation, Zhejiang University.
Zhai, Z., Zhang, X., Jin, Y. (2014). Demographic Consequences of an Immediate Transition to a Universal Two-child Policy. *Population Research*, 38(2).

Figure 1. Locations of Shanxi and Yicheng



Data source: Gibbs Lab at Cornell University.

Figure 2. Crude Birth Rate in Yicheng and other Counties in Shanxi



Data source: Please refer to Section 4.

Figure 3. Difference-in-Difference Analysis

Note: Year 1984 is taken as the baseline year.

Figure 4. Crude Birth Rate in Yicheng and "synthetic Yicheng"

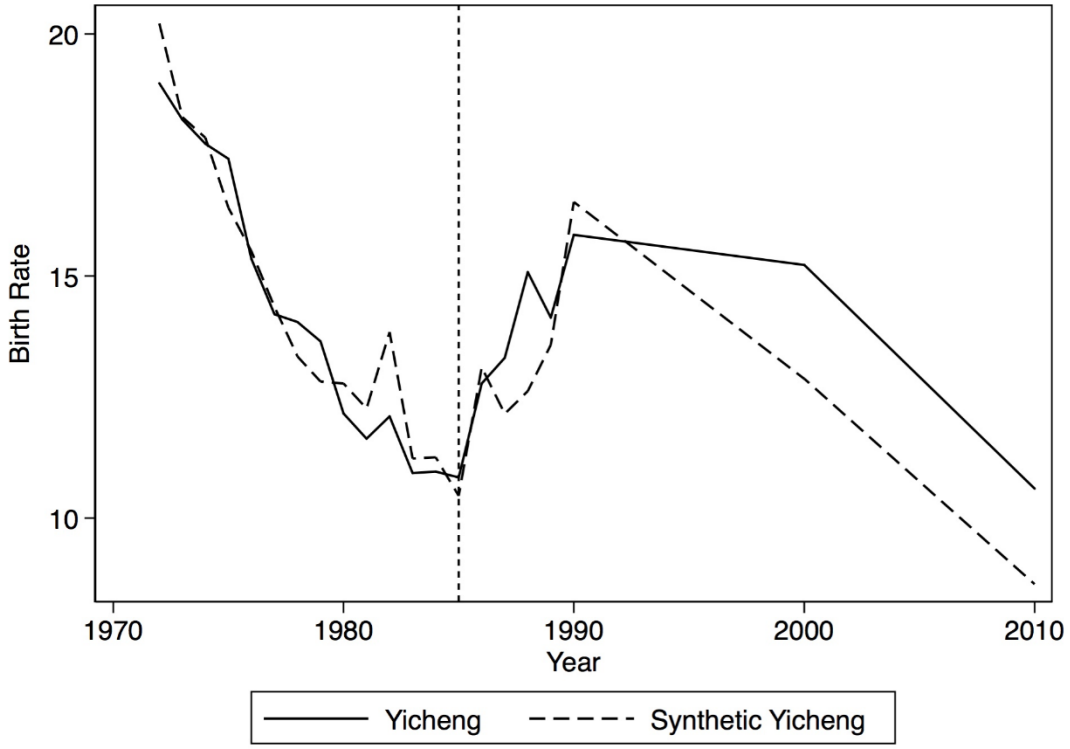
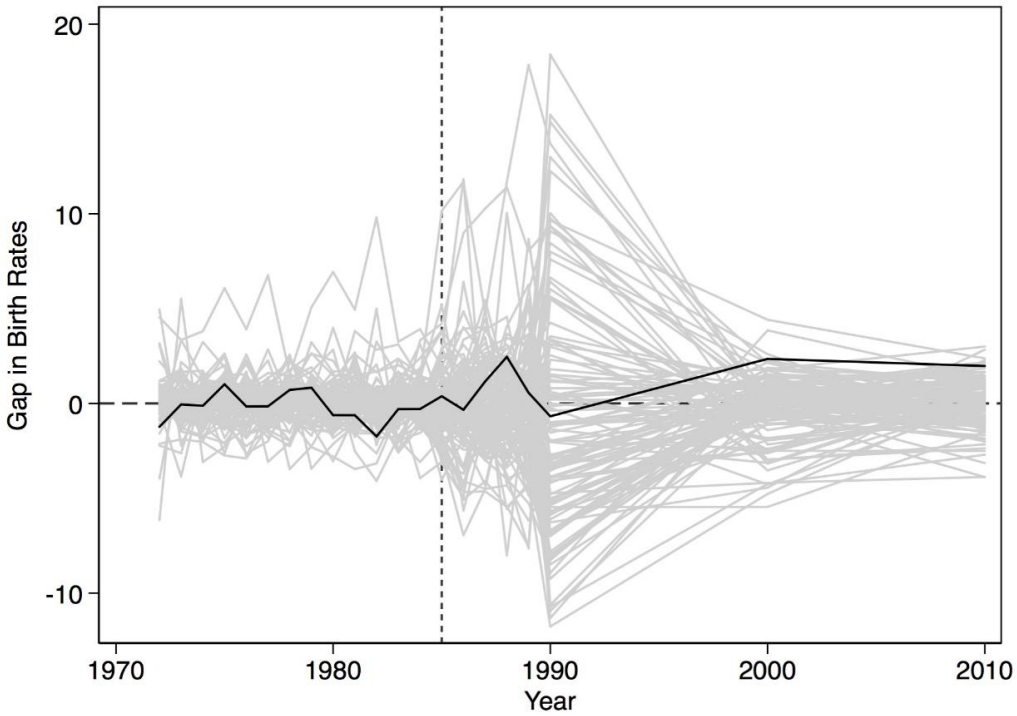


Figure 5. Placebo tests

(a) All 115 counties



(b) Yicheng and 61 counties with smaller pre-treatment MSPE than Yicheng

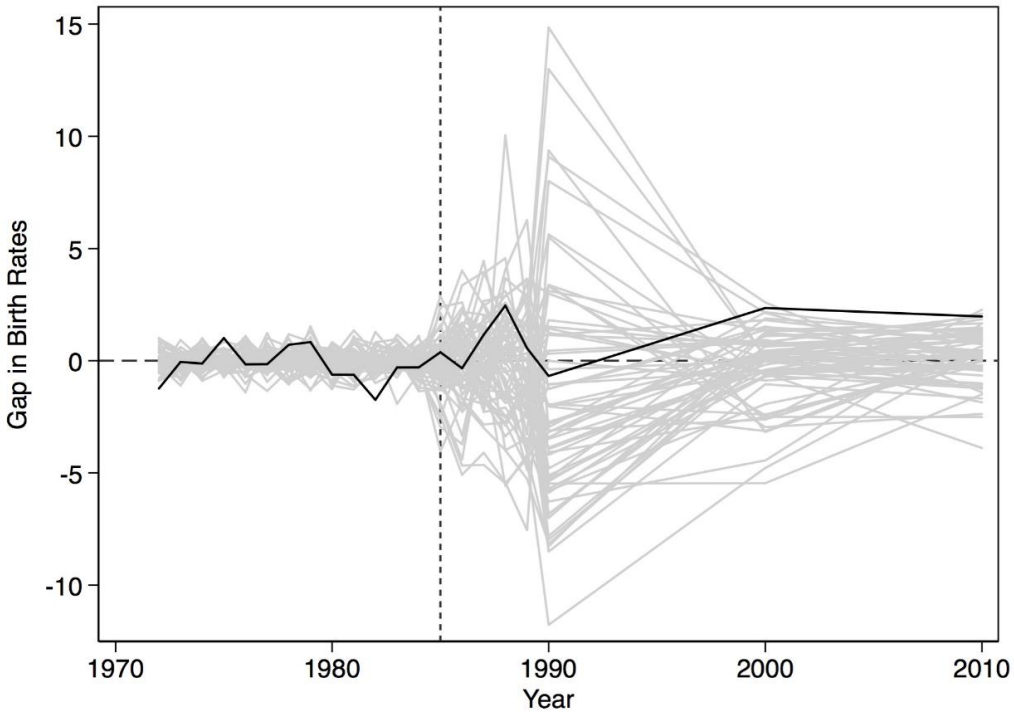


Figure 6. Policy effects based on census samples

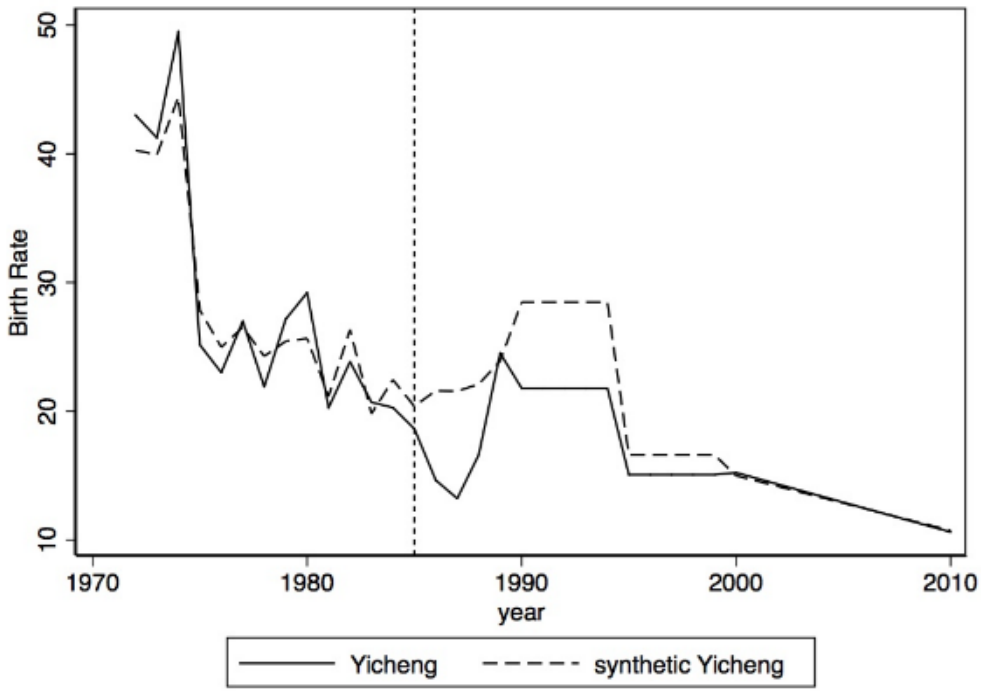
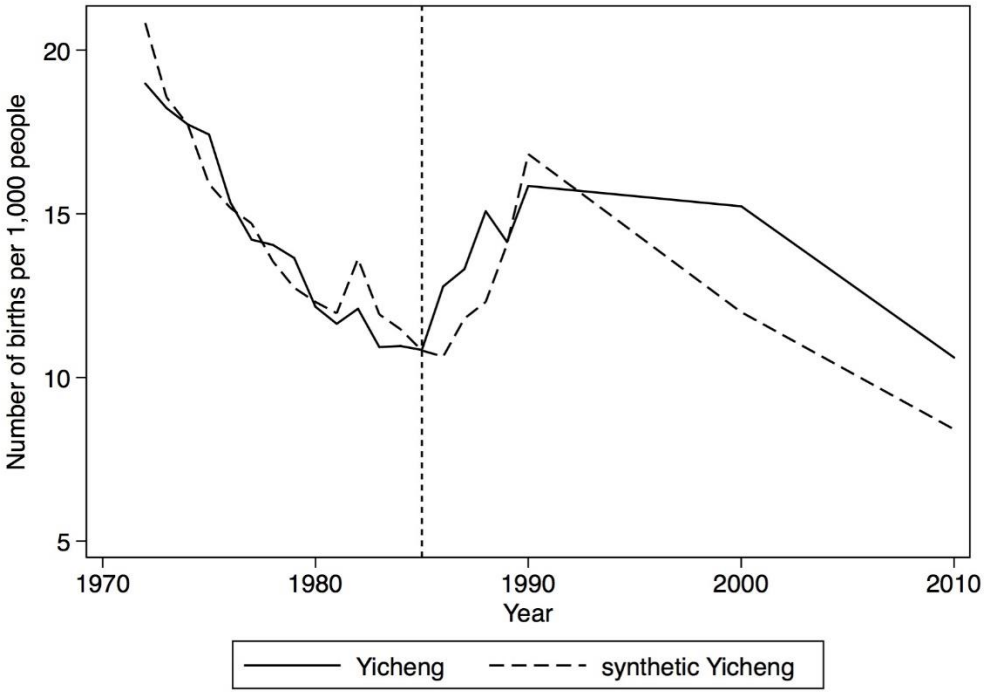


Figure 7. Robustness Check

(a) Using a small donor pool surrounding Yicheng



(b) Extending the pre-treatment period to 1970

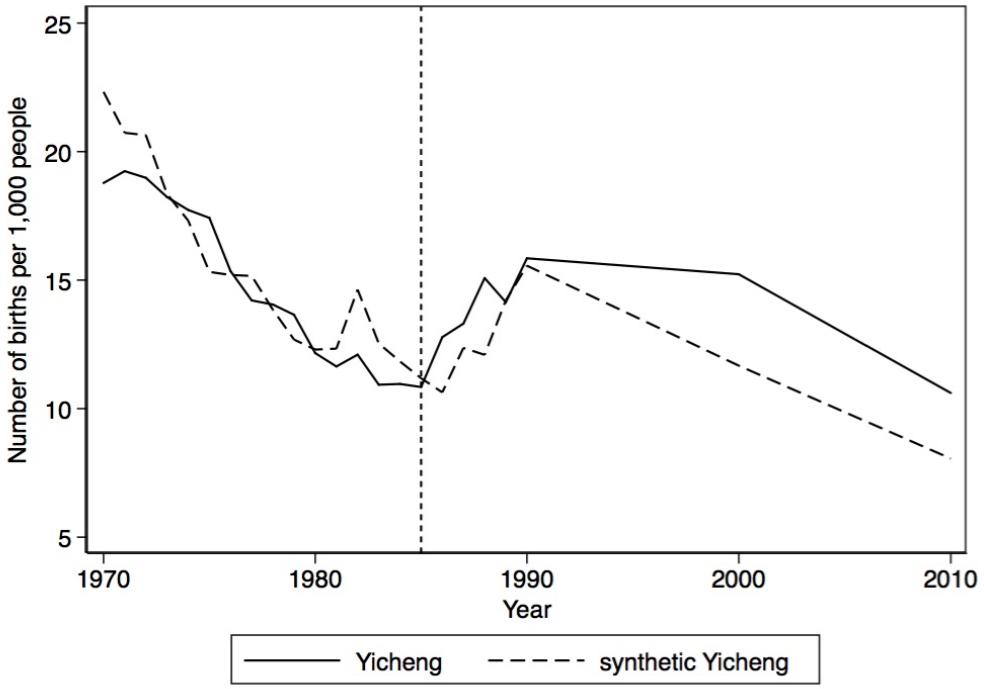


Table 1. Weights of Synthetic Control Groups

County	Weight
Kuangqu, Datong	0.302
Hexi, Taiyuan	0.283
Fushan, Linfen	0.195
Qinyuan, Changzhi	0.184
Pianguan, Xinzhou	0.036

Table 2. Balancing Properties

	Yicheng	Synthetic Yicheng	Average of control counties
% Female, 0-14 (city level)	17.05	15.02	16.12
% Male, 0-14 (city level)	17.98	16.34	17.43
% Female, 15-59 (city level)	28.32	29.15	27.90
% Male, 15-59 (city level)	29.60	33.16	30.73
% Female, 60+ (city level)	3.58	3.08	3.84
% Female (county level)	48.60	42.79	47.16
Deaths per 1,000 people	7.28	5.28	7.12
Hospitals per 1,000 people	0.07	0.22	0.18
% Agricultural population	92.98	56.26	84.48
Log of GDP per capita	5.24	6.17	5.86
Log of rural personal income	4.38	4.54	4.56
Elementary schools per 1,000 people	2.24	1.78	1.75
Middle schools per 1,000 people	0.40	0.41	0.44
Births per 1,000 people (1972)	18.98	20.22	29.70
Births per 1,000 people (1973)	18.23	18.28	27.40
Births per 1,000 people (1974)	17.73	17.85	26.39
Births per 1,000 people (1975)	17.42	16.41	23.67
Births per 1,000 people (1976)	15.35	15.51	19.76
Births per 1,000 people (1977)	14.21	14.36	17.58
Births per 1,000 people (1978)	14.05	13.34	15.91
Births per 1,000 people (1979)	13.65	12.82	15.99
Births per 1,000 people (1980)	12.16	12.78	16.16
Births per 1,000 people (1981)	11.64	12.26	17.01
Births per 1,000 people (1982)	12.10	13.84	18.78
Births per 1,000 people (1983)	10.93	11.23	15.23
Births per 1,000 people (1984)	10.96	11.26	14.42

Data source: please refer to Section 4.

Table 3. Percentage of counties having smaller treatment effects than Yicheng

Year	Full sample	Restricted sample
1985	69.6	74.2
1986	43.5	38.7
1987	68.7	74.2
1988	89.6	85.5
1989	61.7	61.3
1990	58.3	62.9
2000	96.5	96.8
2010	92.2	95.2
# counties	115	62