Stanford | King Center on Global Development

The Limits and Consequences of Population Policy: Evidence from China's Wan Xi Shao Campaign

Kimberly Singer Babiarz

Paul Ma

Grant Miller

Shige Song

October, 2020

Working Paper No. 2000

NBER WORKING PAPER SERIES

THE LIMITS AND CONSEQUENCES OF POPULATION POLICY: EVIDENCE FROM CHINA'S WAN XI SHAO CAMPAIGN

Kimberly Singer Babiarz
Paul Ma
Grant Miller
Shige Song

Working Paper 25130
http://www.nber.org/papers/w25130

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2018, Revised October 2020

This paper was previously circulated under the titles "Limits (and Human Costs) of Population Policy: Fertility Decline and Sex Selection Under Mao", and "Early Fertility Decline and Sex Imbalance in Rural China." We thank Jay Bhattacharya, David Canning, Daniel Goodkind, Giovanna Merli, Nancy Qian, Christine Valente, and Jennifer Van Hook for helpful comments and suggestions. We also thank seminar participants at Columbia, Duke, Hitotsubashi, Penn, UC Berkeley, UC San Diego, University of Michigan, USC, UT Austin, and SIEPR. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Limits and Consequences of Population Policy: Evidence from China's Wan Xi Shao Campaign
Kimberly Singer Babiarz, Paul Ma, Grant Miller, and Shige Song
NBER Working Paper No. 25130
October 2018, Revised October 2020
JEL No. J1,J12,J13,J18

ABSTRACT

Prior to the famous One Child Policy, China's total fertility rate declined by more than 50% during the 1970s - one of the most rapid sustained fertility declines documented in modern history. Coinciding with this transition was China's first national population policy, Wan Xi Shao, also known as the Longer, Later Fewer (LLF) campaign. Studying LLF's contribution to fertility and fertility strategies favoring sons, we find that the campaign i) reduced China's total fertility rate by 0.88 births per woman (explaining 27% of China's modern fertility decline), ii) doubled the use of male-biased fertility stopping rules, and iii) promoted postnatal selection (implying 200,000 previously unrecognized missing girls). Considering Chinese population policy to be extreme in global experience, our paper demonstrates the limits of population policy in explaining demographic transitions— and its potential human costs.

Kimberly Singer Babiarz
CHP/PCOR
117 Encina Commons
Stanford, CA 94305
babiarz@stanford.edu

Paul Ma
Carlson School of Management
University of Minnesota
321 19th Ave S.
Minneapolis, MN 55455
paulma@umn.edu

Grant Miller
Center for Health Policy/
Center for Primary Care & Outcomes Research
Stanford University
615 Crothers Way
Stanford, CA 94305-6006
and NBER
ngmiller@stanford.edu

Shige Song
Queens College of the City
University of New York
Department of Sociology
65-30 Kissena Blvd
Queens, NY 11367
Shige.Song@qc.cuny.edu

A data appendix is available at http://www.nber.org/data-appendix/w25130

1 Introduction

Historically, many governments have used fertility policies to achieve social, political, and economic goals. Policies to reduce or even forcefully restrict fertility rates have their roots in classical macroeconomic growth theory (Keynes and Volcker, 1920; Coale and Hoover, 1958) and a more recent economics literature establishing that larger households reduce living standards and exacerbate income inequality (Chu and Koo, 1990; Galor and Weil, 2000; Hazan and Berdugo, 2002; Kremer and Chen, 2002; Moav, 2005). Alternatively, countries facing below replacement-level fertility rates have introduced pronatalist policies in an attempt to address the macroeconomic consequences of population 'graying' (Milligan, 2005; Bloom et al., 2011; Raute, 2019).

However, the extent to which population policy is ultimately able to influence fertility behavior, particularly in lower-income countries, has been fiercely debated (Pritchett, 1994; Schultz, 1994; Lee, 2003). These debates span several broad areas, including i) the credibility of population policy and family planning program effect estimates (Miller and Babiarz, 2016) and ii) the efficacy of such policies relative to demand-based determinants of ferility (e.g., rising levels of education among women (Barro and Lee, 1993) or increases in the opportunity cost of time (Becker, 1960; Schultz, 1973). Moreover, a parallel concern has been the negative unintended consequences of such policies (e.g., Ebenstein (2010); Howden and Zhou (2014); Jayachandran (2017)). In no country have these concerns been more salient than in China, whose population policies are the most restrictive - and strictly enforced - in global experience.

In this paper, we shed new light on these issues by examining the consequences of China's first national population policy, *Wan Xi Shao*, which coincided with a demographic transition in China that ranks as one of the fastest in global history (Figure 1, Panel A) (Banister, 1987; Chen, 1984; Feeney and Wang, 1993). Literally meaning "Later, Longer, Fewer," this policy (henceforth "LLF") aimed to limit fertility by promoting marriage at older ages ("Later"), longer birth intervals ("Longer"), and fewer lifetime births ("Fewer").

Strikingly, particularly given that China accounted for one-quarter of the world's population when the policy was implemented, China's total fertility rate (TFR) fell by more than 50%, from approximately 6 births per women to 2.75 as the LLF policy was scaled-up (Figure 1, Panel B).[1] Relying on the time-series evidence, many demographers and policymakers have attributed the majority of this decline to LLF (Bongaarts and Greenhalgh, 1985; Lavely and Freedman, 1990; Feeney and Wang, 1993; Bhrolcháin and Dyson, 2007; Goodkind, 2017).[2] By contrast, the more famous One Child Policy was enacted when fertility rates were already approaching replacement level (Cai, 2010). Given its stringent population policies along with its rapid demographic transition and economic development, China's *Wan Xi Shao* policy likely represents an upper bound on the size of feasible family planning program effects.[3]

To the best of our knowledge, this paper provides the first direct evidence on the LLF policy's contribution to fertility decline in China - and its potentially unintended consequences.[4] Digitizing archival records of LLF implementation and matching them to individual-level survey data measuring fertility behavior over several decades, we study behavioral responses along each targeted behavioral margin (age at marriage, birth intervals, number of births). We first establish the logic and validity of an event study framework exploiting the program's staggered implementation across provinces (as early as 1970 and as late as 1979), and importantly, we show that its implementation appears unrelated to either pre-existing trends in fertility or changes in the underlying demand for children (Preston et al., 1978; Rosenzweig and Schultz, 1983; Becker, 1991). We then combine econometric and

[1] The Total Fertility Rate (TFR) is a summary index of age-specific fertility rates for a given period of time (a given year, for example). It is, therefore, a 'period' rate, measuring the expected number of lifetime births that a woman experiencing each age-specific fertility rate in that period would have.

[2] Notably, many of the demand-side determinants of global fertility decline (wage growth, changes in the opportunity cost of time, etc.) were not changing rapidly in China during these years.

[3] While family planning programs in most countries are voluntary, focusing on reducing the costs of fertility control to minimize unwanted births (Glasier et al., 2006), China imposed explicit fertility limits which were strictly enforced for decades (Greenhalgh and Li, 1993; Scharping, 2013; Greenhalgh and Winckler, 2005; Mosher, 2008; White, 2006).

[4] We note that Goodkind (2017) makes cross-national comparisons between China and other countries, drawing inferences about the combined consequences of the LLF policy and the One Child Policy. Since writing our initial draft, two recently released working papers study the role of the Family Planning Leading Groups in China's 1970s fertility decline (Chen and Huang, 2018) and subsequent intergenerational effects (Chen and Fang, 2018).

demographic methods in a novel way to estimate the policy's overall contribution to China's demographic transition.

On fertility behavior, we find that overall, the program reduced China's TFR by about 0.88 births per woman, accounting for about 27% of China's fertility decline prior to 1980 (implying 15.8 million averted births). To investigate the extent to which this TFR decline simply reflects delayed childbearing (e.g., Bassi and Rasul (2017)) rather than reductions in lifetime fertility, we decompose TFR changes into 'quantum' (number of births) and 'tempo' (birth timing) effects (Bongaarts and Feeney, 1998). Doing so, we find that although women's average age at marriage increased by nearly a year, quantum effects account for over 95% of the TFR decline associated with the LLF policy — meaning that TFR changes are overwhelmingly the result of fewer lifetime births.[5] These results suggest that even China is not a marked outlier in global experience with family planning programs (Miller and Babiarz, 2016).

We then turn to the possibility of unintended consequences of the LLF policy - in particular, its potential contribution to fertility strategies reflecting son preference (hereafter 'sex composition strategies'), including sex selection. Son preference is a well-established phenomenon in China, and theory predicts that when there is son preference, fertility decline should promote sex composition strategies favoring male births (Das Gupta and Mari Bhat, 1997; Li et al., 2000; Jayachandran, 2017; Anukriti, 2018).[6] Before prenatal ultrasound technology was commonly available in China, as in our case,[7] there were two sex composition strategies that couples could use: (1) male-biased fertility stopping rules (the practice of having children until reaching the desired number of sons - hereafter "stopping rules"), or

[5]Examining birth intervals directly, we also find little evidence that LLF influenced birth spacing.

[6]Son preference may be rooted cultural practices such as patrilineal marriage (married couples live with the husband's family, making sons critical for economic security in old age) and also be reinforced by economic incentives embedded in village life (in China's communal system, fewer work points were awarded for women than men, etc.) (Arnold and Zhaoxiang, 1986; Coale and Banister, 1994; Ebenstein and Leung, 2010; Ebenstein, 2014; Greenhalgh and Li, 1993; White, 2006).

[7]The introduction and rapid spread of ultrasound technology during the early 1980s was largely coincident with the One Child Policy circa 1980 (Chen et al., 2013). Numerous studies of the One Child Policy find that sex ratios at birth rose rapidly during the 1980s, largely through prenatal ultrasound screening and sex-selective abortion (Almond et al., 2017; Banister, 1987; Chen et al., 2013; Ebenstein, 2011).

(2) postnatal sex selection (hereafter "postnatal selection") through relative underinvestment in girls — and in the extreme, female infanticide.[8] There is aggregate time-series evidence that the LLF policy coincided with increasing use of sex composition strategies in China — both through greater use of stopping rules (Arnold and Zhaoxiang, 1986), which do not alter population sex ratios, and through postnatal sex selection (at third and higher parity births among couples without a son — see Figure 2), producing male-biased sex ratios (Coale and Banister, 1994; Babiarz et al., 2019).[9] If the LLF policy did in fact contribute to population sex imbalance, this would represent an important unintended consequence of population policy.

We develop a novel empirical approach for distinguishing between the use of stopping rules and postnatal selection when prenatal selection is not technologically feasible. Our approach relies on two key facts. The first is that both stopping rule use and postnatal selection increase the probability that couples discontinue childbearing after the birth of a boy (Yamaguchi, 1989), enabling us to estimate the prevalence of *any* sex composition strategies. The second is that only postnatal selection leads to male-biased sex ratios (when prenatal selection is not feasible). Using these facts, we find that the LLF policy increased the use of both strategies, but about 91% of incremental new use of these strategies due to the policy was the use of stopping rules. Specifically, the share of couples using fertility stopping rules rose from 3.25% to 6.8%, while the share of couples practicing postnatal sex selection rose from nil prior to the LLF policy to 0.29% (implying that 0.3% of all births involved postnatal selection by the late 1970s). Although small in relative terms, this prevalence of postnatal selection implies about 200,000 additional "missing girls" in China directly attributable to the LLF policy, explaining about 21% of all girls missing from Chinese birth cohorts during the 1970s. Moreover, because postnatal selection overwhelmingly occurred during the first year

[8] Infanticide was a practice well-documented historically (King, 2014; Lu and Mungello, 2010; Wolf and Huang, 1980) and a concern raised by governmental officials in policy deliberations (Greenhalgh and Winckler, 2005)

[9] Throughout this paper, we define the sex ratio at birth in the conventional way: the number of male live births for each 100 female live births. Research suggests that the biologically 'expected' ratio of male to female births is approximately 105-106 males for every 100 females (Johansson and Nygren, 1991).

of life (and is not generally explained by misreporting — a concern we consider at length in Section 3 and in the Online Appendix), infanticide in particular may have been an important unintended consequence of the LLF policy.

Our paper makes contributions to several broad literatures. First, in contrast to some of the prevailing wisdom in the demography literature, it demonstrates that even one of the most intensive family planning programs in global history (LLF) explains a relatively small share of China's rapid demographic transition (Pritchett, 1994; Miller and Babiarz, 2016; Wang et al., 2016) - and by extension, suggests that the repeal of One Child Policy may do little to increase long-run fertility in China (Wang et al., 2016). In doing so, it also contributes to a better understanding of one of the most rapid demographic transitions on record (Banister, 1987). Second, our estimation strategy is the first to combine a credible observational study design with classical demographic methods (Van Hook and Altman, 2013) - specifically, generating econometric event-study inputs required by the canonical life table population accounting framework - a contribution that we believe has wide applicability in future research on economic demography. Third, although a number of studies have shown that fertility policies have induced prenatal sex selection, to the best of our knowledge, ours is the first to demonstrate that such policies can also lead to *postnatal selection* - a dramatic and perverse unintended consequence, even in the "missing women" literature [citations]. [10]

The paper proceeds as follows. Section 2 provides background on population policy in China, and Section 3 describes our data. Section 4 then presents methods and results for fertility behavior, and Section 5 presents our model, empirical methods, and results for sex composition strategies. Section 6 concludes.

[10] In a recent paper studying fertility and sex selection, Almond et al. (2017) show that the staggered roll-out of land reform in 1979 induced sex selection behavior at second parity among households with a first-born daughter independent of the contemporaneous rollout of the One Child Policy (but under the constraints imposed by it).

2 Background and Context

2.1 Fertility Decline during the Mao Era and the *Wan Xi Shao* ("Later, Longer, Fewer") Policy

At the time of the communist revolution, China's TFR was high, hovering around 6 births per woman in rural areas.[11] Figure 1 shows that it remained stable at this level throughout the 1950s until the Great Leap Famine (1959-1961), when it dropped precipitously to about 3, then rebounded rapidly to pre-famine levels, and again remained relatively constant at about 6 through the end of the 1960s. After the famine, the Chinese government began considering 'management' of population growth to be a critical part of economic planning (Greenhalgh and Winckler, 2005; White, 2006).[12] This perspective was formalized in LLF policy beginning in 1970, and the central government officially codified formal birth planning targets in its Third Five Year Plan (1971-1975).[13]

The LLF policy was a set of broad central government regulations to be designed in detail and implemented by China's provincial and local governments (Greenhalgh, 2008). In practice, provinces implemented the policy between 1970 and 1979 (as Figure 3 shows). Although the historical record does not offer an account of this variation in implementation timing, the degree to which it was related to underlying changes in the demand for children is an important issue that we explore in detail in Sections 4.1 and 4.2. Overall, LLF sought to reduce crude birth rates in rural areas to 15 per 1,000 population through three primary

[11] During the years we study, approximately 85% of China's population lived in rural areas.

[12] Limiting population growth was considered integral to China's economic development and the prosperity of its planned economy. During the LLF and One Child Policy periods, technocrats set birth planning targets in coordination with economic development goals — for example, population growth models were used together with grain production forecasts to set long term birth planning goals (Song et al., 1985). Savings to the state were calculated according to the provisions required for each averted birth avoided (3 million averted births in Anhui province were calculated to save 900,000 tons of grain and 1.6 million meters of cotton cloth, for example) (Zheng et al., 1981). For a detailed account of the political evolution of birth planning in China, see Greenhalgh (2008); Greenhalgh and Winckler (2005); White (2006).

[13] There were two earlier birth planning campaigns (1954-1958 and 1961-1966) that were small and focused on urban areas, featuring fewer restrictions and weaker enforcement (Wang, 2012; Scharping, 2013).

mechanisms: (1) Later marriage — delaying marriage to ages 23 and 25 (for rural women and men, respectively); (2) Longer birth intervals — increasing birth intervals to a minimum of four years; and (3) Fewer lifetime births — limiting couples to 2-3 children in total (Greenhalgh, 2008; White, 2006).[14]

To implement the LLF policy, provincial leaders established birth planning offices and mid-level coordinating committees, which translated central government guidelines into provincial- and local-level targets and managed the daily activities of local birth planning cadres. At the grass-roots level, barefoot doctors, birth attendants, and maternal health aids served as birth planning officers charged with ensuring that births in their localities did not exceed quotas.[15] Specific responsibilities included deciding which couples would receive permission to have a child, delivering free oral contraceptives to couples' homes, tracking which couples had intrauterine devices (IUDs), and persuading couples to undergo sterilization.[16] The birth planning workforce recruited to enforce LLF was vast: in Sichuan province, for example, historians suggest that there was a birth planning officer for every 100 persons — or approximately one million birth planners in total.

Although LLF was technically a voluntary program, birth quotas were taken very seriously (Whyte et al., 2015). On the supply-side, local-level cadres had strong career incentives to meet their targets, leading to political commendation, which was critical for career advancement. On the demand-side, cadres were also allowed to create strong incentives for compliance among households. Compliant households received paid rest periods, higher wages, better housing, and larger staple allocations, for example (Greenhalgh and

[14] The most common forms of birth control during this period were intrauterine devices (50%), sterilizations (25%), and oral contraceptives (8.5%). However, abortions were also common methods of avoiding unplanned births, with an estimated 5 million abortions performed per year during the 1970s (Jowett, 1986).

[15] Commune- and brigade-level birth allowances were determined using a 'top-down-bottom-up' process of negotiation in which targets proposed at higher levels were adjusted according to feedback from grassroots birth planning cadres with knowledge of local fertility demand (Greenhalgh, 2008; Freedman et al., 1988).

[16] According to some accounts, health workers so closely managed fertility in their jurisdictions that they monitored the menstrual cycles of all fertile-age women, posting menstrual cycle status in public forums, and may have even performed monthly exams of women to verify birth control compliance (White, 2006). Other reports suggest that when the number of eligible couples was substantially greater than the number of births permitted under a quota, birth planning cadres would force couples to negotiate among themselves which would be allowed conceive in a given year (Chen and Kols, 1982).

Winckler, 2005). Alternatively, birth officers could increase work assignments, administer public condemnation, or restrict food rations, medical care, and other public services as punishment for failure to comply (Greenhalgh and Li, 1993).[17] Couples were subjected to intense pressure to comply, and historians document many reports of coercion and abuse, including reports of threats and multi-day sessions in which couples were berated until they agreed to abortions (White, 2006; Whyte et al., 2015).

2.2 Population Sex Imbalance

Rooted in patrilineal traditions, a large body of research documents a strong preference for sons in China — and male-biased population sex ratios throughout China's history (Das Gupta and Shuzhuo, 1999; Ebenstein, 2014; Ebenstein and Leung, 2010; Greenhalgh and Li, 1993; Jayachandran, 2015). Historical accounts of China's Imperial Period report the practice of female infanticide as early as the third century BC (Lee, 1981).[18] By the end of the Imperial era, in the late 19th century, some scholars suggest that 10-25% of all newborn girls across all social strata were victims of infanticide (King, 2014; Lee and Wang, 1999). In more recent history, sex ratios were abnormally high during years of famine and political turmoil early in the 20th century due to infant abandonment, infanticide, and differential neglect of girls during childhood (Banister, 1987; Greene and Merrick, 2005; King, 2014; Langer, 1974; Lee and Wang, 1999; Lu and Mungello, 2010; Wolf and Huang, 1980).[19]

The vast majority of research on population sex imbalance in modern China focuses on the One Child Policy, land reform, and the coincident diffusion of ultrasound technology

[17] Anecdotal reports also describe public meetings to publicly criticize and shame non-compliant individuals — a powerful political tactic used throughout the Cultural Revolution (White, 2006). Other reports describe couples being forced to attend day-long meetings where they were subjected to intense pressure to have abortions (including late-term abortions) or to undergo sterilization.

[18] Philosophers writing in the third century B.C. and historical legal texts provide describe female infanticide early in China's history (Lee, 1981; Jimmerson, 1990). There are also accounts of infanticide being practiced as a form of 'birth control' in the 11th century (Ebrey, 1993) and 17th century (Mungello, 2008).

[19] The ratio of men to women born during the 1920s and 1930s appears to have ranged between 107.3 and 113.6, peaking during the 1940s at 112.7-117.7. Although it is not possible to discern if these imbalanced sex ratios emerged at birth or reflect differential mortality throughout childhood and early adulthood, qualitative records suggest that much of this imbalance began at birth (Song, 2012).

across the country during the 1980s and later. These changes led directly to the phenomenon of sex-selective abortion, which became widespread, resulting in a dramatic rise in sex ratios at birth among cohorts born in the 1980s and more recently (Banister, 1987; Chen et al., 2013; Gupta, 2005; Ebenstein, 2010, 2014; Ebenstein and Leung, 2010; Almond et al., 2017; Hull, 1990; Yi et al., 1993). However, because theory predicts that fertility decline should lead to sex selection in a population preferring sons (Das Gupta and Mari Bhat, 1997; Das Gupta and Shuzhuo, 1999; Jayachandran, 2017; Jayachandran and Kuziemko, 2011), there is reason to suspect that sex selection and population sex imbalance may have emerged during China's rapid fertility decline throughout the 1970s - prior to the One Child Policy.

Figure 2 shows that sex selection behavior may in fact have risen during the 1970s (earlier than generally recognized). Among couples presumably having the greatest demand for sons (those having children at third or higher parity — and not yet having a boy), sex ratios at birth actually rose as high as 115-121 by the end of the 1970s (Babiarz et al., 2019). Because this increase in sex ratios at birth occurred before ultrasound technology was generally available, it could also suggest a resurgence of infant abandonment or infanticide. Scaling the sex ratios at birth in Figure 2 by the size of China's population during the 1970s, these ratios imply over 950,000 additional missing girls in China (Babiarz et al., 2019).[20]

3 Data and Measurement

For our empirical analyses of fertility decline and sex composition strategies, we use data from three major types of sources: (1) Archival public health records (*Weishengzhi*) and provincial annals from 28 Chinese provinces; (2) Individual-level fertility history records from China's 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception; and (3) Province-year economic and demographic data from both China's official provincial yearbooks

[20] See Babiarz et al. (2019), and the Online Appendix for detailed calculations.

and the China Family Panel Survey.[21] We describe each data source below (Table 1 shows descriptive statistics).

3.1 Data Sources

First, we obtained provincial LLF policy implementation dates from provincial public health archives (*Weishengzhi*) and historical provincial annals. These records document public health campaigns and other provincial government activities in each province and year from the 1950s through the 1990s. Official provincial committees published these records and statistics using data from epidemiological surveillance stations, provincial health department archives, local government registers, and other administrative sources (Babiarz et al., 2015). We interpret the first mention of birth planning regulation in each province (specifically, age at marriage, birth spacing, and overall fertility) in these archival records to signify implementation of the LLF policy.[22] Figure 3 shows LLF policy implementation years in each Chinese provinces, and details are provided in Appendix Table A1.

Second, we use retrospective fertility history records from China's 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception. This nationally representative survey of married women ages 15 and older includes 400,000 married women living in rural areas.[23] A key feature of this survey is that it collected complete fertility histories from each woman interviewed (akin to the World Fertility Survey and its successor, the Demographic

[21]These 28 provinces are Anhui, Beijing, Fujian, Gansu, Guangdong, Guangxi, Guizhou, Hainan, Hebei, Heilongjiang, Henan, Hubei, Hunan, Inner Mongolia, Jiangsu, Jiangxi, Jilin, Liaoning, Ningxia, Qinghai, Shaanxi, Shandong, Shanghai, Shanxi, Sichuan, Tianjin, Xinjiang, and Yunnan.

[22]In cases in which exact birth planning regulation dates were not explicitly reported, we generally use the date on which provincial Birth Planning Leadership Committees were established (see Appendix Table A1 for details). Our approach is conservative in adopting the earliest possible date that birth planning program activities may have begun (birth planning committees may have taken several years to fully scale-up policy implementation). To explore the quality of our policy timing measurement, we conduct a placebo test, randomly re-assigning province-level implementation years across our sample in each of 1,000 iterations. Appendix A8 plots both the resulting empirical distributions and our program effect estimates. In general, our estimates lie outside of the traditional confidence intervals of these empirical distribution.

[23]Because the LLF policy was implemented differently in rural and urban areas (marriage age targets and the number of children allowed varied across rural and urban settings, for example) (White, 2006; Lavely and Freedman, 1990), and because the overwhelming majority of births (87%) in our sample occurred in rural areas, we restrict our analysis to rural households. Appendix Figure A6 shows that our results are robust to the inclusion of urban residents.

and Health Surveys), recording the timing and location of all births and deaths of respondents' children back to the 1960s — yielding a sample of more than 1.2 million births during our study period between 1964 and 1979. Although this sample is not strictly representative back in time, it nonetheless permits internally valid estimation.[24] In contrast to most research on population history in China, which uses population census data, we also highlight that the "Two-Per-Thousand" survey enables us to study fertility decline and sex composition strategies among population subgroups expected to have greater demand for sons (higher parity couples not yet having a boy, for example).[25]

Third, we use province-year information from a variety of sources to account for other factors that may have influenced fertility and sex-selection during our study period. These include measures of economic development, the child mortality environment, and social instability associated with China's Cultural Revolution (1966-1976). Specifically, we obtain data on provincial Gross Domestic Product (GDP), total grain output, primary school enrollment rates, and rural population share from China's official provincial yearbooks, which are produced by the National Bureau of Statistics of China (and supplemented by the China Statistical Data Compilation (China Statistical Bureau, 2000)).[26] We calculate child mortality rates for a given province and year as the rate at which children under age 5 die as reported in our survey, averaged over the preceding 5 years.[27] To capture the influence of other major programs likely to influence child survival (programs providing vaccinations, mosquito and other vector control services, sanitation, and basic nutrition), we also digitize records from China's provincial public health archives (*Weishengzhi*). Finally, using the China Family Panel Survey, we measure the intensity of the Cultural Revolution (and associated social

[24]Because the survey is representative of married women in 1988, selective mortality among women presumably means that it is not strictly representative of women/births in earlier years (Gakidou and King, 2006).

[25]Although our data do not contain complete migration history information, migration in China was restricted and highly controlled under China's household registration system during our study period. Only 11% of mothers in our survey lived in a province other than their province of birth, and the vast majority of those women migrated prior to marriage.

[26]Some of this data is available from the University of Michigan's China Data Center: http://chinadataonline.org

[27]Although no reliable source of child mortality data is available, we test the sensitivity of our results to using alternative sources of mortality data — for example, vital statistics records of overall mortality.

instability) by calculating the share of people in each birth cohort and province who were 'sent-down' — a practice of sending college-age youth away from their homes to work on collective farms during the Cultural Revolution (Institute of Social Science, 2014).

Table 1 shows summary statistics for variables used in our analysis.

3.2 Measurement of "Missing Girls"

Because we consider deviations from the naturally occurring sex ratio at birth (prior to the introduction of ultrasound technology) to reflect unreported girls that died early in life (i.e., postnatal selection), a note about data quality is warranted.[28] Under-reporting of births and under-enumeration of living children (and living girls in particular) during the 1980s and 1990s is well documented in the demography literature (Cai and Lavely, 2003; Goodkind, 2011; Merli and Raftery, 2000; Yi, 1996; Zhang and Zhao, 2006). However, existing literature suggests that the degree of such under-reporting during the 1970s was substantially less (Coale, 1984; Coale and Banister, 1994; Yi, 1996).

To the best of our knowledge, however, no previous work has directly assessed the degree of under-reporting during the 1970s in the 1988 "Two-Per-Thousand" survey — including under-reporting by birth order and under-reporting of girls relative to boys. Given the importance of this concern for the interpretation of our empirical results, we use three methods to investigate the extent to which unreported girls lived beyond infancy as unregistered children in our sample.

3.2.1 Comparing Sex Ratios using the 1982 and 1990 Population Censuses

First, following Coale and Banister (1994), we directly investigate the extent to which possibly unreported female births in the 1988 "Two-Per-Thousand" survey 're-appear' as adult women in China's population censuses, focusing on those births most likely to be

[28] While under-reporting may also have been possible through the adoption of girls, the design of the 1988 survey explicitly differentiates between adopted and non-adopted children. Enumerators were instructed to ensure that pregnancy histories only reflected own children (including those subsequently given up for adoption) and excluding children who are adopted (Babiarz et al., 2019).

underreported. We use one percent microsamples of the 1982 and 1990 Chinese population censuses to compare sex ratios at birth (number of male births for each 100 female births) for each birth cohort reported in the 1988 fertility survey with sex ratios for the same birth cohorts as reflected in the 1982 and 1990 censuses, adjusting for differential mortality using reverse survival methods. We make these comparisons among all births, and births by parity and sibship sex composition to rule out the possibility that higher parity births may be more likely to be underreported.[29]

3.2.2 Comparing Implied Population Counts to Population Census Data

Second, following Coale (1991) we use the "Two-Per-Thousand" survey to calculate the age-specific rate at which women deliver male and female babies in each year. We then apply these fertility rates by maternal age and child sex (simultaneously) to age-specific population counts of women reported in population census microsamples (interpolated between the 1964 and 1982 censuses), yielding an estimate of the total number of boys and girls born in each calendar year. We then compare the estimated number of male and female births implied by these calculations to the actual number of individuals in each birth cohort reflected in the 1982 and 1990 censuses to estimate the degree of underreporting for boys and girls by birth cohort in the fertility survey.[30]

3.2.3 Comparisons with the 1982 "One-Per-Thousand" Fertility Survey

Third, we compare the 1988 "Two-Per-Thousand" national fertility survey directly to the 1982 "One-Per-Thousand" survey (which is generally considered good quality — but that has important limitations) (Banister, 2004; Bhrolcháin and Dyson, 2007; Coale and Banister, 1994). For every woman surveyed in the "One-Per-Thousand" survey, we identify woman surveyed in the "Two-Per-Thousand" with exactly the same characteristics. Pooling these matched observations we then investigate whether the year in which a woman was surveyed

[29] See the Online Appendix, Appendix Figures A1-A3, and Appendix Tables A2-A3 for details.

[30] See Online Appendix, and Appendix Figure A4 for details.

predicts her total number of births reported, the sex ratio of her children, or the number of sons/daughters reported.[31]

Overall, all three approaches suggest little systematic differential underreporting of girls (and importantly, little under-reporting by parity and sex composition of previous births) in the 1988 survey. The Online Appendix presents each of these methods and corresponding results in detail.

4 Fertility Behavior

We begin by studying how couples' fertility behavior changed in response to the LLF policy, focusing on each behavioral margin targeted by the policy: age at marriage (*Wan*, or "Later"), birth intervals (*Xi*, or "Longer"), and completed lifetime fertility (*Shao*, or "Fewer"). Before doing so, however, we first assess the identifying assumptions underlying many of the econometric and demographic methods that we subsequently use. In assessing these assumptions, we also establish whether or not there is any *prima facie* evidence of a fertility response to the implementation of the LLF policy.

A brief note about the methods we use throughout the paper is also warranted. To the extent possible, we use an event study framework to analyze how distinct dimensions of fertility and behavior reflecting son-preference change in response to the LLF fertility control policy. In some cases, however, other frameworks (or modifications) are appropriate — for example, when modeling the duration of some outcomes for which policy rules vary by age (in the case of marriage and birth timing) or when cell sizes otherwise become prohibitively small (for sex selection behavior among population subgroups). We note and explain these cases as they arise.

[31]See Online Appendix for details.

4.1 Identifying Assumptions

Given that the introduction of the LLF policy across provinces was not randomly assigned, a concern is that provinces may have implemented the policy in response to underlying changes or trends in the demand for children in each province (a concern about which historical accounts are largely silent). However, we highlight that Chinese government planners generally lacked incentives to respond to the preferences of provincial residents, and Figure 4 provides *prima facie* evidence consistent with this view (White, 2006). Controlling for province and year fixed effects, it plots important determinants of the demand for children by event year (normalizing the year of LLF implementation in each province to be event year zero).[32] Specifically, trends in provincial GDP, the child mortality rate (under age 5), population share working in agriculture, and total provincial grain production are flat, hovering around zero (with reasonable precision) prior to LLF implementation — suggesting that the introduction of LLF in each province was not correlated with changes in these key determinants of demand.

We next evaluate this concern further by directly examining the relationship between policy timing and pre-existing trends in a key basic measure of fertility — the annual risk of parity progression (or probability of birth).

4.2 Parity Progression Estimation

We use an event study framework to estimate the relationship between the introduction of the LLF policy and the annual risk of parity progression, exploiting the staggered introduction of the LLF policy across provinces and over time. We focus on fertility responses among sub-groups expected to have differential behavioral responses to the policy, enabling us to consider our underlying identifying assumptions more extensively.

[32] For research on the importance of these factors for the demand for children, see Preston et al. (1978); Rosenzweig and Schultz (1983); Schultz (1985); Becker (1991).

Specifically, we estimate variants of the following equation using woman-year observations among those under age 40:

$$\begin{aligned} Birth_{ijy} &= \alpha \\ &+ \phi_{Son} + \lambda_{Parity} + \rho_{EventYear} \\ &+ \mu_{Son \times Parity} + \eta_{Son \times EventYear} + \xi_{Parity \times EventYear} \\ &+ \psi_{Son \times Parity \times EventYear} \\ &+ X_i\beta + Z_{jy}\theta + \delta_j + \gamma_y + \epsilon_{ijy} \end{aligned} \tag{1}$$

where $Birth_{ijy}$ is an indicator for whether or not mother i in province j delivered a child in year y. We use a linear probability model to regress this outcome on a set of indicator variables for whether or not mother i already has at least one surviving son, indicators for maternal parity (0, 1, 2, 3, and 4+), time in years between year y, and the year of LLF implementation in province j ('event year,' ranging from -8 to +8) along with all two- and three-way interactions.[33] We also control for maternal and household characteristics X_i (a mother's highest level of education, her age at marriage, and the ethnicity of her household head) as well as province-year characteristics Z_{jy} (provincial GDP, a five-year moving average of the under-5 mortality rate, gross agricultural output, grain production, and the proportion of the population classified as rural). Finally, Equation 1 also includes provincial fixed effects (δ_j) and calendar year fixed effects (γ_y), absorbing unobserved time-invariant provincial attributes and changes over time common across China, respectively. Because our number of clusters is relatively small, we construct confidence intervals by wild cluster bootstrap with 1,000 replications (Cameron et al., 2008).[34]

[33] We define indicators for the birth parity at which a mother is at risk, from 1^{st} birth through 4^{th} and higher parity births (grouping higher parity births together). Because we focus on how behavioral responses change over time and across parity, we use linear probability models for ease of interaction term interpretation (Ai and Norton, 2003), but logit models yield similar predicted birth hazards for each parity and event year group (results available upon request).

[34] Throughout the paper, wherever possible, we assess statistical significance using the wild cluster bootstrap method. However, this approach is not well-suited for several analyses, including life table calculations, and

Figure 5 shows estimates from Equation 1, with results for each parity shown in separate panels (and separate plots for couples with and without a son at second and higher parity within each panel). Consistent with our identifying assumptions (and Figure 4), at all parities there is no clear pattern of pre-existing fertility trends systematically related to the timing of LLF implementation, suggesting that the introduction of the policy across provinces was also unrelated to past fertility behavior (as well as underlying trends in the demand for children).[35]

However, we find abrupt changes in parity progression following the implementation of the LLF policy. At first parity, the probability of a birth actually rises in the later years of the LLF policy — a result that may suggest shorter birth intervals following postponed marriages in the late LLF period (because we consider only married women to be at risk for a first birth), as shown in Section 4.4.[36] The probability of a second birth then declines modestly for parents both with and without a son following the implementation of policy, with a somewhat steeper decline (although not significantly so) among couples with a son. Consistent with a true program effect, this decline then becomes more pronounced at third and at fourth and higher parities. Moreover, at third parity (the general fertility limit imposed by the policy — and therefore the parity at which the policy has the most 'bite'), the gradient is significantly steeper among couples with a son — among whom the demand for more children should be weaker. Overall, the pattern of reductions by parity and whether or not a couple already has a son is consistent with a true effect of the LLF policy on fertility behavior.[37]

Finally, to explore the sensitivity/robustness of our results to alternative specifications, we re-estimate variants of Equation 1 1) using provincial time trends, 2) excluding groups of

cross-specification prevalence rate calculations. As an alternative, we instead use a pairs-cluster bootstrap method for Figures 6 and 7 as noted.

[35] No point estimates for parity 1-2 births are statistically significant prior to the introduction of LLF. Some pre-LLF point estimates among 3rd parity and 4th and higher parity births are statistically different from zero, but we do not find any discernible trends preceding the introduction of LLF.

[36] Section 4.3 shows that age at marriage rises under LLF, and Section 4.4 shows that the interval between marriage and first birth declines.

[37] At fourth parity, the decline is again steeper among couples with a son, but not significantly so — presumably because the penalties for violating the policy were already imposed after third parity births. Detailed results in tabular form available upon request

control variables, 3) adding urban couples to the sample, and 4) excluding each province one-by-one. Appendix Figures A5-A7 show that our estimates are robust in all cases. Additionally, to assess the quality of our data on program implementation dates, we conduct a placebo test, drawing 1,000 sets of randomly re-assigned provincial implementation dates and plotting the resulting empirical distributions for each parity and sibship sex composition group. Appendix Figure A8 shows these results. In general, the estimates using our program implementation dates fall outside of the 0.05% tail of these empirical distributions (with estimates for second and third parity couples without previous sons having the largest p-values, 0.055 and 0.059, respectively).

4.3 Age at Marriage (***Wan***)

We next study changes along the first behavioral margin targeted by LLF — age at marriage (Wan). To model duration until marriage, we use a woman-year sample including each woman from age 15 until marriage to estimate discrete-time hazard models of the following general form:[38]

$$\begin{aligned} Marriage_{ijy} &= \alpha \\ &+ \phi_{LLF} + \lambda_{age} + \rho_{age \times LLF} \\ &+ X_i\beta + Z_{jy}\theta + \delta_j + \gamma_y + \epsilon_{ijy} \end{aligned} \tag{2}$$

where $Marriage$ is a dummy variable for whether or not woman i in province j marries in year y, ϕ is an indicator for whether or not LLF was active in province j and year y, λ is a vector of dummy variables for women's ages a, ρ is a vector of interactions between the LLF policy indicator and each age dummy, and all other variables as defined before.[39] Note that

[38] Because divorce rates were very low in China during this period (the crude divorce rate, or the number of divorces per 1000 population in a given year, was approximately 0.3 in 1978 (Dommaraju and Jones, 2011)), we simplify our analysis by studying only the age at first marriage.

[39] A discrete-time hazard model is more appropriate than other duration models such as a Cox proportional hazard model because it does not require an assumption about constant proportional hazards over time. A commonly-cited advantage of Cox proportional hazard models is that they address censoring of duration

because we estimate interactions between the LLF policy and dummy variables for single years of age, sample sizes in event year by age cells become too small to use an event study framework. Instead, we use a single policy indicator variable, capturing the average program effect.

Using a logit specification, the estimated odds ratio $exp(\phi_{LLF})$ captures the effect of LLF on the probability of marriage at the reference age (age 23 — the marriage age generally mandated by LLF) among those not yet married. For each age a, the coefficients $\rho_{a \times LLF}$ then reflect changes in this program effect at all other ages 15-40 relative to age 23. Because Ai and Norton (2003) show that the standard marginal effect calculation for nonlinear models is incorrect for interaction terms, we instead adopt a prediction-based approach for obtaining age-specific marginal effects of the LLF policy. Specifically, we first use estimates from Equation 2 to predict the likelihood of marriage at each age, both with and without LLF, holding all control variables constant at observed pre-LLF values. We then interpret the difference between these predicted marriage hazards at each age as the marginal effect of the LLF policy on age-specific probabilities of marriage among those not yet married (Buis et al., 2010). We compute confidence intervals using the pairs-cluster bootstrap method. [40]

Figure 6 Panel A shows the marginal effect of the LLF policy on age-specific probabilities of marriage among women. We find that the probability of marriage falls at ages below age 23, with age-specific estimates that are statistically different from 0 between ages 17 and 21.[41] Alternatively, at ages 23 and higher, the probability of marriage rises and is statistically significant (becoming less precise at the oldest ages, where there is little mass in the distribution of marriage ages). This pattern of results reflects fewer marriages before age

variables, but we note that by construction, our sample does not contain censored observations (only married individuals were surveyed). We also note that because we use a sample of ever-married women, we are only able to study realized age at marriage.

[40]Because our approach to estimating age-specific marginal effects relies on post-estimation predicted marriage probabilities, we bootstrap the distribution of coefficients using a cluster bootstrap rather than bootstrapping individual Wald statistics (the wild bootstrap).

[41]Prior to the introduction of LLF, the minimum age of marriage for women was 18 years old under the 1950 Marriage Law (Kane, 1987). However, our data show that the mean age at marriage prior to the LLF policy was 19-20 years of age.

23, the mandated minimum marriage age under the LLF policy, and more marriages at older ages.

To measure the implied change in age at marriage due to the policy, we use a single decrement life table approach to map estimated changes in age-specific marriage hazards q_x to predicted changes in the distribution of age at marriage (Van Hook and Altman, 2013). Specifically, we predict survival curves l_x describing the share of women remaining unmarried at each age $15 \leq x \leq 40$ both with and without the LLF policy. Beginning at age 15, at which all women enter the risk set, we calculate the share of women remaining at risk at each age after 15 using estimated age-specific marriage hazards $l_x = l_{x-1} - (l_{x-1} \times q_{x-1})$. Figure 6 Panel B plots the inverse of these survival curves (i.e., cumulative density functions (CDFs)). Consistent with Panel A, there is a statistically significant shift to the right in the distribution of age at marriage under LLF, implying an increase in the median age at marriage of 8.2 months [95% CI: 4.3-11.2 months]. Appendix Table A4 provides full life table results.

4.4 Birth Intervals (*Xi*)

We next estimate how birth intervals changed under the LLF policy (*Xi*, the second targeted behavioral margin). Following the same approach for estimating changes in age at marriage, we first study changes in age at first birth, re-estimating Equation 2 — but with woman-year observations from age 15 to first birth and using a dummy variable for whether or not woman i in province j has her first birth in year y as the dependent variable.

Figure 7 Panel A plots marginal effects of the LLF policy on the likelihood of a first birth at each age. The age pattern of changes in first births closely tracks that of marriage, with reductions in risk of first birth before age 23 and increases at older ages. Figure 7 Panel B shows corresponding cumulative density functions with and without LLF (predicted in the same way as age at marriage), with a statistically significant increase in the median age at first birth of 5.2 months under LLF [95% CI: 3.5-6.8 months]. Overall, these results suggest that first births generally followed closely after marriage, both before and after LLF.[42]

We then also estimate how subsequent birth intervals (marriage to first birth and intervals at higher parities) changed with LLF. Restructuring our sample as woman-quarter observations beginning three quarters after either marriage or a previous birth (the approximate gestational period) and ending at the next birth, we otherwise use a similar approach.[43] Specifically, stratifying by parity, we estimate discrete hazard models of the following general form for parities 1-4:

$$\begin{aligned} Birth_{ijq} &= \alpha \\ &+ \phi_{LLF} + \lambda_{Quarter} + \rho_{Quarter \times LLF} \\ &+ X_i\beta + Z_{jy}\theta + \delta_j + \gamma_y + \epsilon_{ijq} \end{aligned} \tag{3}$$

where $Birth$ is an indicator variable for whether mother i in province j advances in parity q quarters after her previous birth and all other variables are defined as before.

Figure 8 shows cumulative density functions implied by these results for births at each parity (constructed in the same way as the CDFs for age at marriage). Although the results are potentially suggestive of small changes in birth intervals, they are not statistically significant for any interval.[44]

[42] Appendix Table A4 shows these results.

[43] To isolate the policy effects on the timing of births from the effects on lifetime fertility, the sample is restricted to mothers who eventually advance in parity at some point prior to the survey enumeration.

[44] Our results imply that on average, the interval between marriage and first birth increases by 1.6 months, and subsequent birth intervals increase by 2.4, 1.4, and 1.1 months (none statistically distinguishable from zero). However, comparing later years of the LLF policy (5 or more years after implementation) to pre-LLF years, we find that the median interval between marriage and first birth declined by 2.7 months. This result is roughly equivalent to the decline implied by changes in age at marriage and age at first birth — and consistent with the pattern of first parity progression estimates shown in figure 5, panel A. Results available upon request.

4.5 Fertility (*Shao*)

A natural approach to estimating completed fertility effects of LLF would be to compare the lifetime births of women at all fertile ages when LLF was implemented with the lifetime births of women who were menopausal when LLF was introduced. In practice, however, two problems prevent us from adopting this approach. First, we do not observe completed fertility among all women in our sample (some of whom were still fertile at the time of the 1988 survey). Second, because the One Child Policy was introduced in 1980, births between 1980 and 1988 reflect the influence of the One Child Policy (and not just LLF).

We therefore develop an alternative approach, building on methods developed by Van Hook and Altman (2013) and used in Section 4.3. First, we use a discrete-time hazard model to estimate the inputs needed to build sequential multi-decrement life tables for each LLF event year. These life tables then yield corresponding total fertility rates (TFRs) that are conditional on the same covariates that we use to study other dimensions of fertility behavior, isolating variation in the TFR driven by the LLF policy.[45] However, TFRs are period measures that summarize Age-Specific Fertility Rates (ASFRs) at a given point in time, and TFR changes under LLF reflect both 'quantum' (number) and 'tempo' (timing) fertility effects. Because our interest is completed fertility (quantum effects), we then decompose the TFR change due to the LLF policy into its separate quantum and tempo components following Bongaarts and Feeney (1998). Isolating the quantum effect of the LLF policy effectively enables us to recover estimates of changes in completed lifetime fertility.[46]

[45] This approach implicitly assumes that households believed that the LLF birth planning policy would be permanent. If households believed that LLF was temporary and would reverse within their fertility window, compliance may have been higher than it would have been under the belief that the policy would be permanent, causing us to over estimate the policy effect. On the other hand, if households believed that the policy would become even more strict (e.g., One Child Policy), compliance might have been lower than it would have been under the belief that the policy would be permanent, leading to an underestimation of lifetime fertility effects. While we are not aware of any data on household expectations from the period, the tightening trend in each successive fertility control campaign, combined with increasing birth planning budgets (Scharping, 2013) suggest that the latter scenario is more likely.

[46] Another approach would be to restrict our sample to women who were at least age 40 by the time of the One Child Policy (and thus could reasonably be assumed to be unaffected by it). However, this approach

4.5.1 Total Fertility Rate (TFR) Estimation

Using a discrete-time hazard framework similar to Equation 1, we use logit models to estimate:

$$\begin{aligned} Birth_{ijy} &= \alpha \\ &+ \phi_{EventYear} + \lambda_{AgeGroup} \\ &+ \rho_{EventYear \times AgeGroup} \\ &+ X_i\beta + Z_{jy}\theta + \delta_j + \gamma_y + \epsilon_{ijy} \end{aligned} \tag{4}$$

where λ is a vector of maternal age group indicators (age 15-19, 20-24, 25-29, 30-34, 35-39, 40-45), and all other variables are defined as before. To ensure that mothers' parity is unique within each five-year age interval, we stratify by parity, estimating separate models for women at risk of parity 1-7 births. Women enter each parity-specific sub-sample either at age 15 or after a birth at the previous parity, and they exit the sub-sample (progressing to the next) either at the time of their next birth or at age 45.

Following Van Hook and Altman (2013), we then use estimates from Equation 4 to predict conditional birth hazards by maternal age and parity for each LLF event year. Because we allow event year indicators ($\phi_{EventYear}$) and their interactions with maternal age groups ($\rho_{EventYear \times AgeGroup}$) to vary, holding all other covariates constant at values observed in the year prior to the LLF policy, the discrete change in predicted birth hazards can be interpreted as the marginal effect of LLF on age- and parity-specific fertility.

As Appendix Table A5 shows, these birth hazards (q_x) form the first part of each event year life table, and we then use them to calculate both corresponding survivor functions (l_x) and age- and parity-specific birth rates (d_x) (Appendix Table A5 describes these calculations in detail). Summation across the d_x tables' rows yields ASFRs, and summation down the

would not allow for the effect of LLF on fertility at younger ages because the sample would be restricted to women 30 years old and above in 1970, around the time of the first LLF program initiation. Furthermore, our data show that only a small proportion of births occur between ages 30 and 40.

columns yields parity-specific fertility rates. Summation again across either of the ASFRs or the parity progression ratios yields the TFR for a given event year (at bottom-right, as shown in Appendix Table A5).[47]

Table 2 shows the resulting regression-adjusted ASFRs and TFRs by event year. Relative to the year of LLF implementation, the TFR decline due to the policy was about 0.9 births, explaining about 27% of China's overall TFR decline during these years.[48] To see more clearly the age- and parity-specific fertility changes underlying this TFR effect, Figure 9 graphically depicts these changes by age and parity during the 8 years following implementation of the LLF policy. The greatest reductions occurred at third parity among women in their late 20s as well as at higher parities among women in their 30s. Figure 10 then summarizes the overall effect of the program on China's TFR over time, showing the unadjusted Total Fertility Rate observed in each event year and the counterfactual TFR without the LLF policy implied by our estimates.

4.5.2 Fertility Quantum and Tempo Decomposition

To then decompose this change in TFR into its quantum (number) and tempo (timing) components, we compute tempo-adjusted Total Fertility Rates (TFR') for each event year (Bongaarts and Feeney, 1998).[49] This is simply the sum of parity-specific fertility rates in each event year, adjusted for the mean change in age at childbirth at each parity:

$$TFR' = \sum_{p} \frac{TFR_p}{1 - r_p} \tag{5}$$

[47] Our focus is estimating changes in TFRs due to the LLF policy (rather than recreating observed fertility rates). It is important to note that our TFRs should differ from observed TFRs for three reasons. First, we estimate life tables for event years rather than calendar years. Second, our life tables hold all control variables constant at our sample means, which are averages across both pre- and post-LLF years. Third, because births above parity 7 were very rare, we estimate age- and parity-specific fertility rates up to parity 7 (but omit higher parity births).

[48] From the earliest year of the LLF policy (1970) to the start of the One Child Policy in 1979, China's TFR fell by 6-2.75=3.25 births (Wilmoth et al., 2007). Our estimates therefore suggest that LLF was responsible for a TFR decline of 0.88 births, or 0.88/3.25=27.1%.

[49] Although Kohler and Philipov (2001) discuss the importance of variance effects in this decomposition, the yearly change in the variance of age of childbearing before vs. after the LLF policy is only 1.18%.

where TFR_p is the parity p-specific fertility rate for a given event year and r_p is the change in mean age at childbirth (in months) at each parity in that same event year. We obtain estimates of r_p directly from Section 4.3 (assuming that age at first birth increased at a constant rate over time).

Table 2, column 7 shows the resulting tempo-adjusted TFR' for each event year.[50] Consistent with our finding of little change in birth intervals under the LLF policy, the tempo-adjusted change in quantum fertility between event year 0 and 8 is close to the overall change in TFR — 0.82 births, accounting for 94% of the decline in the overall change in TFR due to LLF.

5 Son Preference-Based Fertility Strategies

Given the relationship that we find between the LLF policy and fertility behavior — and past research establishing a positive correlation between fertility decline and sex selection, we next examine how the fertility strategies couples used to achieve their desired family composition (sex composition strategies) changed in response to the policy. In our environment, there are two strategies that couples desiring a boy could use. The first is simply to have children until obtaining the desired number of boys — that is, to use a male-biased fertility stopping rule (Clark, 2000; Jensen, 2003; Yamaguchi, 1989). The second is postnatal sex selection — either preferential treatment of sons over daughters (leading to relatively higher mortality rates among daughters than otherwise expected), or in the extreme, female infanticide.

Conceptually, as fertility costs rise (due to a restrictive population policy, for example), the use of both stopping rules and postnatal selection could increase. This is because we consider the LLF policy to have two effects. (1) First, both delayed age of marriage and increased birth spacing raise the opportunity cost of all children, which decreases the number of

[50] Appendix Table A6 shows the complete set of results.

children couples would like to have (hereafter, 'target family size'),[51] even absent a preference for sons. (2) Second, the LLF fertility limit increases the marginal cost of children beyond the target in a nonlinear way. The first implies that families are less likely to have sons by chance, conditional on their new target family size, a phenomenon known in demography as "sex selection pressure" (Li et al., 2000). The second implies more postnatal selection among families with sufficiently low psychological costs relative to the cost of exceeding target fertility.

Because existing frameworks do not examine these issues formally, the next section provides a simple model to illustrate them. In doing so, it formalizes several well-established empirical results in demography that characterize demographic phenomena when there is a preference for sons — for example, girls will have more siblings than boys (Clark, 2000; Jensen, 2003; Basu and De Jong, 2010), and couples' terminal births are more likely to be boys (Yamaguchi, 1989; Park and Cho, 1995). An implication of our model (that these strategies will generally be used, or be reflected, at the terminal birth) also directly guides our empirical approach to distinguishing stopping rule use and postnatal selection.

5.1 Model

Following Ben-Porath and Welch (1976) and Jayachandran and Kuziemko (2011), our model assumes that risk-neutral couples desire a target number of children (i), and because of a preference for sons, face a trade-off between their desire for a minimum number of sons and the total number of children that they have. The following utility function captures these preferences:

$$u(s, n) = -\theta(n - i)^2 + \lambda \ln(s + 1) \tag{6}$$

[51] We assume that households have a target number of children they would like to have taking the full cost of having and raising children into account — a target which is independent of their desire for sons. Importantly, we distinguish this 'target number of children' from standard demographic measures such as the 'ideal number of children', 'desired total fertility,' and 'wanted total fertility' (Pritchett, 1994).

The first term represents an inverted u-shaped preference over the total number of children, and the second term captures a couple's preference for sons. Absent son preference, utility is maximized when the realized number of children, n, is equal to the target number of children i, which is determined by both demand- (e.g., the opportunity cost of children) and supply-side factors (LLF and the costs of fertility control generally). θ represents the disutility incurred from deviating from target fertility (e.g., penalties from violating LLF targets). In the second term, $\lambda \geq 0$ is a parameter for the intensity of son preference (the utility a couple experiences from having s sons independent of n). When a couple prefers sons over daughters ($\lambda > 0$), and when couples reach fertility size $n = i$ without any sons, a natural tension arises as these couples face a trade-off between the disutility from exceeding the target family size ($n > i$) and the marginal expected utility of a birth through a potential son.

These preferences embed the demography literature on stopping rule behavior as a sex composition strategy (Clark, 2000). For example, a couple desiring at least one son may adopt the strategy of having up to 2 children, regardless of sex, and proceeding with a third terminal birth only if the first two are female. Numerically, this strategy is rationalized with the model preferences of $i = 2, \theta = 1$ and $\lambda = 3$ in Equation 6.[52] With the use of stopping rules, couples that already have at least one son will choose to stop childbearing at parity i regardless of the sex of the parity i child.[53] Among couples without a son, the sex of the parity i birth will determine if the couple also chooses to have a parity $i + 1$ child. As a result, couples will be more likely to discontinue childbearing after a son is born — producing the well-known result that stopping rules increase the probability that the terminal child (i.e., the youngest child) is male (Yamaguchi, 1989), and leading to a pattern whereby, on average, females have more siblings than males (Jensen, 2003).

[52] Under such a strategy, households stopping at parity i must have at least one son, whereas those stopping at a parity above i include families which may not have a son. The average share of sons at the household level is therefore lower for smaller realized family sizes — hence the negative association between the share of sons and family size (Clark, 2000; Jensen, 2003).

[53] Without loss of generality, we assume that the minimum number of sons desired is one, but our model may be generalized to allow for sex composition strategies used to achieve a minimum of two or more sons.

To study how LLF changes couples' choice of sex composition strategies, we augment this basic framework to allow for postnatal selection through the neglect of a newborn child for a cost c (which includes the psychological cost of neglecting just-born children).[54] The choice of a particular sex composition strategy becomes necessary when couples reach their target family size ($n = i$) without achieving the desired number of sons.[55] With this number of total children, and for all realizations in the number of sons s, the expected utility of using a stopping rule but not neglecting the child is:

$$\underbrace{-\theta}_{\text{disutility of additional birth}} + \underbrace{\frac{\lambda}{2}(\ln(s+2) + \ln(s+1))}_{\text{expected utility of additional birth}}$$

and the expected utility of neglecting a child and trying again for a son is:

$$\underbrace{-\frac{1}{2}(\theta + c)}_{\text{expected disutility of additional birth}} + \underbrace{\frac{\lambda}{2}(\ln(s+2) + \ln(s+1))}_{\text{expected utility of additional birth}}$$

Using a stopping rule is therefore preferred to neglect when:

$$\frac{\lambda}{2}(\ln(s+2) - \ln(s+1)) - \theta > 0; c > \theta \tag{7}$$

while neglect is preferred when:

$$\frac{\lambda}{2}(\ln(s+2) - \ln(s+1)) - \frac{1}{2}(\theta + c) > 0; c < \theta \tag{8}$$

The first terms of Equation 7 and 8 reflect the marginal expected benefits of an additional birth while the second terms represent the marginal costs. As θ, the cost shifter of deviating

[54] We note that although not technologically possible during our study period, our framework could be extended to include the cost of prenatal selection (including the financial cost of an ultrasound and abortions as well as associated psychological costs).

[55] It is easy to show that families will naturally have children until ($n = i$) regardless of their sex composition.

from couples' preferred family size increases, the likelihood of neglect in Equation 8 increases relative to the likelihood of stopping rule behavior in Equation 7.

Importantly, households using either sex composition strategy will stop childbearing after having a son (making use of the strategies empirically difficult to disentangle). In the case of postnatal selection (but not the use of stopping rules), however, a greater share of surviving (and recorded) births are male.[56] These two observations directly inform our empirical approach to estimating each strategy.

5.1.1 LLF Predictions

In the absence of son preference, couples simply achieve their target number of children ($n = i$) and then stop childbearing with the terminal child equally likely to be a boy or girl. However, if there is a preference for sons, couples may be willing to exceed their target number of children if necessary to have the desired number of sons. The fertility restrictions imposed by the LLF policy can be considered a supply-side constraint, effectively reducing couples' target fertility. Holding the desired number of sons constant, as the target number of children decreases, the likelihood of having no son (or fewer than the desired number) prior to reaching target fertility increases. As a result, couples must exceed the target family size more often — leading to a higher prevalence of male-biased stopping rule behavior. Simultaneously, because LLF imposes penalties for births beyond the policy limit, deviation from target family size is also more costly under LLF ($\theta' > \theta$) — and hence postnatal selection becomes relatively more attractive as an alternative to continuing with a stopping rule (Equations 7 and 8). In general, our model predicts that sex composition strategies will be used (or become evident) at the terminal birth, a result that again guides our empirical framework below.

In summary, we predict that realized family size will be lower under LLF — and that there will be greater use of both male-biased stopping rules and postnatal selection

[56] We assume that households do not report births resulting in postnatal selection, and that these unreported births are not living as unenumerated children. The Online Appendix tests these assumptions in detail. Furthermore, note that within the model, if a family chooses to selection once, it will continue to do so every time a girl is realized until a boy is eventually born.

(depending on families' psychological costs relative to the cost of exceeding target fertility). The degree to which each increases because of the LLF policy is an empirical question.

5.2 Empirical Estimation of Sex Composition Strategies

Because it is not possible to identify the use of either stopping rules or postnatal selection at the individual or household level (because we do not observe target family size i in Equation 6), we develop an empirical approach for disentangling the two in the aggregate. (1) First, we estimate the prevalence of *any* sex composition strategy (including both stopping rule use and postnatal selection) due to the LLF policy. Because our model predicts that both stopping rules and postnatal selection are used (or become evident) on the terminal birth, we operationalize this measure by estimating excess discontinuation of childbearing following the birth of a boy (relative to the birth of a girl) due to the policy. Absent son preference, the probability of discontinuing childbearing should be unrelated to the sex of the final birth. (2) Second, we directly estimate the prevalence of postnatal selection due to LLF, or the differential likelihood that a given birth is a boy, exploiting the fact that stopping rule use does not change this probability (nor the sex ratio at birth in the aggregate). (3) Finally, subtracting the prevalence of postnatal selection from the prevalence of all sex composition strategies, we recover the prevalence of stopping rule use due to the LLF policy.

5.2.1 Estimating the Prevalence of Any Sex Composition Strategy: Excess Discontinuation of Childbearing after a Boy

Our model predicts that if couples choose a fertility strategy reflecting son preference, using either stopping rules or postnatal selection, they will do so on their terminal birth (or their strategy becomes evident on their terminal birth), which will be the birth of a boy. We therefore consider excess discontinuation of childbearing after a boy (vs. a girl) to be a combined measure of the prevalence of *any* sex composition strategy. Stratifying by parity, we use Ordinary Least Squares to estimate changes in discontinuation after a boy due to LLF using variants of the following equation:

$$\begin{aligned} Stop_{ijy} &= \alpha \\ &+ \phi_{Male} + \lambda_{Period} + \rho_{NoSon} \\ &+ \mu_{Male \times Period} + \eta_{Male \times NoSon} + \xi_{Period \times NoSon} \\ &+ \psi_{Male \times Period \times NoSon} \\ &+ X_i\beta + Z_{jy}\theta + \delta_j + \gamma_y + \epsilon_{ijy} \end{aligned} \tag{9}$$

where $Stop_{ijy}$ is an indicator variable for whether or not the current birth to couple i in province j and year y is the terminal birth, ϕ is an indicator for whether or not the child born is a boy, λ is a vector of dummy variables for period relative to the start of the LLF policy, and ρ is an indicator for whether or not the couple has previously had a son. Note that because sex composition strategy use is infrequent — and hence cell sizes become smaller than in our analyses of fertility behavior, we group third and higher parity births together, and we also group event years into three event periods: years prior to the LLF policy ('pre-LLF,' the omitted group), 1-4 years after LLF implementation ('early LLF'), and 5-8 years after implementation ('late LLF'). Vectors μ, η, ξ, and ψ are two- and three-way interactions between sex, period, and previously born sons, and all other variables are as defined before.[57] Standard errors are estimated using the wild bootstrap method.

Table 3 columns 1-3 show estimates among couples having their first son at each parity (in rows) and in each period (in columns), compared to otherwise similar couples having a girl at the same parity and in the same period. Columns 4-6 then show estimates among couples with at least one previous son.[58] We find that even prior to the LLF policy, couples without sons are more likely to stop childbearing after their first son is born (Table 3, row

[57] We deem a birth to be a couple's "terminal birth" if a minimum of 8 years passes without a subsequent birth.

[58] Table 3 shows the linear combination of coefficients ($\phi_{Male} + \mu_{Male \times Period} + \eta_{Male \times NoSon} + \psi_{Male \times Period \times NoSon}$) (columns 1-3) and ($\phi_{Male} + \mu_{Male \times Period}$) (columns 4-6) for each period estimated from Equation 9.

1) — and increasingly so when the first son is born at higher parities. Specifically, for first through third and higher-parity births, the increase in discontinuation after a boy is 0.5, 8.0, and 13.6 percentage points, respectively (relative to couples having another girl at the same parity in the same period). We also find a more tempered increase in the probability of discontinuation following a son among those with one or more sons (columns 4-6). These results suggest that even prior to LLF, couples prefer at least 1-2 sons (on average) and were using sex composition strategies to have them.

After the LLF policy, use of sex composition strategies generally grows during the early LLF period, and even more so during the late LLF period — and the gradient by parity also persists. Focusing on the late LLF period (5-8 years after implementation of the policy), row 3, column 1 of Table 3 shows that first-time parents having a boy are 8.8 percentage points more likely to stop childbearing relative to those having a girl. Rows 2-3, column 3 then show that among those with no previous sons, parents are 25.7 percentage points more likely to stop childbearing after the birth of their first son at parity 2, and they are 31.0 percentage points more likely when the first son occurs at parity 3 or higher (relative to parents at the same parities having another daughter). Weighting results by the proportion of couples in each parity and sex composition group and by the proportion of couples stopping childbearing at each parity, we find that overall prevalence of any sex composition strategy (i.e., prevalence of discontinuation after a boy) doubled under the LLF policy, rising from 3.3% of couples prior to the policy to 6.8% of couples in the late LLF period (95% CIs: 2.7%-3.8% and 5.5%-7.6%, respectively).[59]

5.2.2 Postnatal Sex Selection

Next, to study the relationship between the LLF policy and postnatal selection directly, we estimate variants of the following equation by Ordinary Least Squares (OLS),

[59]Detailed calculations available upon request.

again stratifying by parity:

$$
\begin{aligned}
Male_{ijy} &= \alpha \\
&+ \lambda_{Period} + \rho_{NoSon} + \xi_{Period \times NoSon} \\
&+ X_i\beta + Z_{jy}\theta + \delta_j + \gamma_y + \epsilon_{ijy}
\end{aligned} \tag{10}
$$

where $Male_{ijy}$ is a dummy variable for whether or not a birth to mother i in province j and year y is a boy, and all other variables are defined as before. For each period, the sum of coefficients $(\rho_{NoSon} + \xi_{Period \times NoSon})$ captures the incremental increase in probability of a male birth among couples with no previous sons (relative to couples of the same parity with at least one previously born son in the same LLF policy period). Because the probability of having a boy (in the absence of prenatal screening technology) should not deviate from the biologically expected rate — unless achieved through postnatal selection, the estimates for $(\rho_{NoSon} + \xi_{Period \times NoSon})$ therefore measure the prevalence of postnatal selection attributable to the LLF policy in each period.

Table 4 shows these results. We do not find evidence of statistically significant postnatal selection prior to the LLF policy, regardless of the sex of the first birth. However, column 2, row 2 shows that early in the LLF period, second parity births were 1.5 percentage points more likely to be a boy when a couple had no prior son (relative to parents with at least one son). Column 3, row 3 then also shows that late in the LLF period, third and higher parity births to couples without prior sons were 2.3 percentage points more likely to be male (relative to parents with at least one son; 95% CI: .0041 - .0425). These results imply a sex ratio at birth of 117 boys per 100 girls among this subset of couples late in the LLF period — and the emergence of postnatal selection in modern China under the LLF policy generally.

At face value, male-biased sex ratios at birth suggest neonatal neglect — or in the extreme, infanticide (we note conflicting qualitative reports about infanticide in China during the 1970s (Banister, 1987)). However, additional selection through neglect of girls can occur later in childhood as well. We explore the possibility of differential child mortality at older ages by re-estimating Equation 10 among children at each year of age one through five; Appendix Table A7 shows these results. Overall, our estimates show that the relationship between the LLF policy and the sex ratios of children reaching ages one through five generally track the results in Table 4 for sex ratios at birth — suggesting that the vast majority of excess female mortality (or selection) may have truly occurred during the first year of life. We also consider the extent to which two alternative explanations may account for our postnatal sex selection results: relatively early weaning of girls (motivated by a desire for another pregnancy to try again for a son) or a "quantity-quality tradeoff" (as male-biased fertility stopping rules increases the number of siblings that girls have relative to boys) (Jayachandran and Kuziemko, 2011; Becker, 1991). Although our ability to test these possibilities is imperfect, Appendix Tables A8-A9 provides evidence using several different approaches suggesting that neither explains the majority of our postnatal selection results.[60]

5.2.3 Male-Biased Fertility Stopping Rules

Finally, we recover the prevalence of stopping rule use due to the LLF policy by subtracting our estimates of postnatal selection (Section 5.2.2) from the share of couples using any sex composition strategy (Section 5.2.1). Table 5 shows these results by parity, previously born sons, and LLF period. Although stopping rule use increased slightly among higher parity couples with at least one previous son, the LLF policy led to the most dramatic increases in stopping rule use among couples with no sons. In rough terms, the use of stopping rules more than doubled under the policy among these couples, rising from 0.3% to 4.6% at first parity, from 4.1% to 13.2% at second parity, and from 7.0% to 13.7% at third and higher parity.

[60] Our calculations imply that the gender gap in breastfeeding could explain 9-33% of missing girls due to the LLF policy.

5.2.4 Sex Composition Strategies by Type

Figure 11 summarizes our results on the use of sex composition strategies, depicting the implied overall prevalence of each strategy over time under the LLF policy (weighting our estimates by the share of couples in each corresponding cell). Overall, Figure 11 shows that the use of stopping rules accounts for the vast majority (91%) of incremental new use of sex composition strategies due to the LLF policy.[61] Specifically, the share of couples using stopping rules approximately doubled under the policy, rising from 3.25% to 6.46% of all couples. Figure 11 also shows the emergence of postnatal sex selection under the policy, with the share of couples using postnatal selection rising from nil to 0.32% of couples. Despite the relatively low rate of postnatal selection, our results nonetheless imply about 200,000 missing girls in China directly attributable to the LLF policy, roughly 21% of the 955,000 missing girls in China during the 1970s (Babiarz et al., 2019).[62]

6 Conclusion

Chinese population policy is widely considered to be a dramatic outlier in the global history of family planning (Robinson and Ross, 2007). Coinciding with a demographic transition that ranks as one of the fastest in global history, beginning in the early 1970s, China established fertility limits and recruited a large network of birth planning workers with broad authority to grant permission for marriages and births, monitor couples' behavior,

[61] To make this calculation, we divide the increase in stopping rule use by the increase in any sex composition strategy: $\frac{3.21}{3.53}$.

[62] To make these calculations, we weight postnatal selection point estimates by the proportion of births occurring in each parity and sex composition group in each period, summing to calculate the proportion of births postnatally selected in each period (as a share of all births occurring in each period): .05% in the early LLF period and about 0.29% in the late LLF period. We then multiply these rates by the total number of births occurring in each LLF period according to vital statistics (China Statistical Bureau, 2000). Our data suggests that approximately 40% of all births during the 1970s fall within the 'early LLF' period (event years 1-4), and 25% occurred 5 or more years after implementation (the 'late LLF' period) — roughly 84.8 and 54.3 million births, respectively. This suggests that 84,815,221 $\times$0.0005023 + 54,247,483 $\times$ 0.002902 = 200,029 missing girls are attributable to the LLF policy (or 200,029 $\div$ 955,000 = 20.9% of all missing births during the 1970s or $\approx$ 0.20% of all female births in the same time period).

and incentivize compliance. Given its intensity and reach, *Wan xi Shao* may provide an upper-bound on the feasible effects of population policy on fertility behavior.

Overall, we find that the Later, Longer, Fewer policy reduced China's total fertility rate by almost one birth per woman, accounting for about 27% of China's overall fertility decline prior to 1980, implying approximately 16.9 million averted births.[63] Decomposing this TFR change into 'quantum' and 'tempo' effects, we show that although the policy raised the median age of first births by 5.2 months, the decline in TFR was largely the result of fewer lifetime births rather than changes in the timing of births.

These results reinforce the view that changes in the underlying demand for children matter most for fertility decline (Pritchett, 1994). Although other scholars have suggested that China's birth planning policies may represent an exception and in fact be the primary force behind its fertility decline (White, 2006; Greenhalgh and Winckler, 2005), our results suggest otherwise. In general, major determinants of the demand for children include economic development, falling infant and child mortality rates (Angeles, 2010; Kalemli-Ozcan, 2002; Schultz, 1985), increasing opportunity costs of women's time (Breierova and Duflo, 2004; Lavy and Zablotsky, 2011; Schultz, 1985), and anticipated increases in future demand for human capital (Galor and Weil, 2000). Many of these forces were not clearly at work in China during the 1970s, however, and we speculate that the exception — declines in infant and child mortality during preceding decades (Banister and Hill, 2004) may have played an important role.

While family planning programs and population policy may have important health and socioeconomic benefits for mothers and their children, including a reduced risk of maternal death (Menken and Rahman, 2001; Cleland et al., 2012; Jain, 2011; Winikoff and Sullivan,

[63] We calculate the approximate number of averted births in the following way. First, we compute the total number of births in China in each event year by weighting the total number of births in each calendar year (China Statistical Bureau, 2000) by the share occurring in each event year (calculating weights using the "Two-Per-Thousand" data). Second, assuming that the percent decline in births in each event year is equivalent to the percent decline in the TFR in the corresponding event year, we compute averted births in each event year by multiplying the percent change in the TFR associated with the LLF policy (see Section 4.5) by the number of births occurring in event year zero. Third, we add averted births across event years, yielding an estimate of 16,908,669 averted births in total.

1987) and both increased human capital investments and lifetime earnings among mothers and children (Canning and Schultz, 2012; Greene and Merrick, 2005; Joshi and Schultz, 2013; Miller, 2010; Pop-Eleches, 2006), our study also shows that there may be human costs as well. Specifically, we develop a new empirical approach for estimating the prevalence of separate sex composition strategies that were technologically feasible in our context, and we show that the LLF policy led directly to an increase in the use of both male-biased fertility stopping rules and postnatal selection (through neglect or possible infanticide). Although postnatal selection was relatively rare, our results imply that the LLF policy resulted in about 200,000 additional missing girls, explaining about 21% of all missing girls during the 1970s. These results are consistent with our model of fertility behavior when couples prefer sons and suggest an important unintended consequence of the LLF policy — and potentially population policy generally — not previously studied.

References

Ai, C. and Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, 80(1):123–129. 16, 19

Almond, D., Li, H., and Zhang, S. (2017). Land reform and sex selection in China. *National Bureau of Economic Research Working Paper*. 3, 5, 9

Angeles, L. (2010). Demographic transitions: Analyzing the effects of mortality on fertility. *Journal of Population Economics*, 23(1):99–120. 36

Anukriti, S. (2018). Financial incentives and the fertility-sex ratio trade-off. *American Economic Journal: Applied Economics*, 10(2):27–57. 3

Arnold, F. and Zhaoxiang, L. (1986). Sex preference, fertility, and family planning in China. *Population and Development Review*, 12(2):221–246. 3, 4

Babiarz, K., Ma, P., Song, S., and Miller, G. (2019). Population sex imbalance in China before the One Child Policy. *Demographic Research*, in press. 4, 9, 12, 35

Babiarz, K. S., Eggleston, K., Miller, G., and Zhang, Q. (2015). An exploration of China's mortality decline under Mao: A provincial analysis, 1950–80. *Population Studies*, 69(1):39–56. 10

Banister, J. (1987). *China's Changing Population*. Stanford: Stanford University Press. 1, 3, 5, 8, 9, 33

Banister, J. (2004). Shortage of girls in China today. *Journal of Population Research*, 21(1):19–45. 13

Banister, J. and Hill, K. (2004). Mortality in China 1964–2000. *Population Studies*, 58(1):55–75. 36

Barro, R. J. and Lee, J.-W. (1993). International comparisons of educational attainment. *Journal of Monetary Economics*, 32(3):363–394. 1

Bassi, V. and Rasul, I. (2017). Persuasion: A Case Study of Papal Influences on Fertility-Related Beliefs and Behavior. *American Economic Journal: Applied Economics*, 9(4):250–302. 3

Basu, D. and De Jong, R. (2010). Son targeting fertility behavior: Some consequences and determinants. *Demography*, 47(2):521–536. 26

Becker, G. (1960). *An Economic Analysis of Fertility. Demographic and Economic Change in Developed Countries*. Princeton University Press. 1

Becker, G. S. (1991). *A Treatise on the Family*. Cambridge, Massachusetts: Harvard University Press. 2, 15, 34

Ben-Porath, Y. and Welch, F. (1976). Do sex preferences really matter? *Quarterly Journal of Economics*, 90(2):285–307. 26

Bhrolcháin, M. N. and Dyson, T. (2007). On causation in demography: Issues and illustrations. *Population and Development Review*, 33(1):1–36. 2, 13

Bloom, D. E., Canning, D., and Fink, G. (2011). Implications of population aging for economic growth. Technical report, Working. 1

Bongaarts, J. and Feeney, G. (1998). On the quantum and tempo of fertility. *Population and Development Review*, 24(2):271–291. 3, 22, 24, 57

Bongaarts, J. and Greenhalgh, S. (1985). An Alternative to the One-Child Policy in China. *Population and Development Review*, 11(4):585–617. 2

Breierova, L. and Duflo, E. (2004). The impact of education on fertility and child mortality: Do fathers really matter less than mothers? *National Bureau of Economic Research Working Paper*. 36

Buis, M. L. et al. (2010). Stata tip 87: Interpretation of interactions in non-linear models. *The Stata Journal*, 10(2):305–308. 19

Cai, Y. (2010). China's Below-Replacement Fertility: Government Policy or Socioeconomic Development? *Population and Development Review*, 36(3):419–440. 2

Cai, Y. and Lavely, W. (2003). China's missing girls: Numerical estimates and effects on population growth. *China Review*, 3(2):13–29. 12

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427. 16, 49, 58, 59

Canning, D. and Schultz, T. P. (2012). The economic consequences of reproductive health and family planning. *Lancet*, 380(9837):165–171. 37

Chen, L. and Huang, Y. (2018). The power of the government: China's family planning leading group and the fertility decline since 1970. *Global Labor Organization Discussion Paper Series 204*. 2

Chen, P.-c. and Kols, A. (1982). Population and birth planning in the peoples republic of China. *Population Reports. Series J: Family Planning Programs*, (25):577–619. 7

Chen, S. (1984). *Fertility of Women During the 42 Year Period from 1940 to 1981.* Beijing: CPIC. 1

Chen, Y. and Fang, H. (2018). The long-term consequences of having fewer children in old age: Evidence from china's "later, longer, fewer" campaign". *National Bureau of Economic Research Working Paper*. 2

Chen, Y., Li, H., and Meng, L. (2013). Prenatal sex selection and missing girls in China: Evidence from the diffusion of diagnostic ultrasound. *Journal of Human Resources*, 48(1):36–70. 3, 9

China Statistical Bureau (2000). *China Statistical Yearbook*. Beijing: China Statistical Bureau. 11, 35, 36

Chu, C. Y. C. and Koo, H.-W. (1990). Intergenerational Income-Group Mobility and Differential Fertility. *The American Economic Review*, 80(5):1125–1138. 1

Clark, S. (2000). Son preference and sex composition of children: Evidence from India. *Demography*, 37(1):95–108. 25, 26, 27

Cleland, J., Conde-Agudelo, A., Peterson, H., Ross, J., and Tsui, A. (2012). Contraception and health. *Lancet*, 380(9837):149–156. 36

Coale, A. J. (1984). *Rapid Population Change in China, 1952-1982*. Washington, DC: National Academy Press. 12

Coale, A. J. (1991). Excess female mortality and the balance of the sexes in the population: An estimate of the number of "missing females". *Population and Development Review*, 17(3):517–523. 13

Coale, A. J. and Banister, J. (1994). Five decades of missing females in China. *Demography*, 31(3):459–479. 3, 4, 12, 13

Coale, A. J. and Hoover, E. M. (1958). Population growth in low income countries: A case study of indiaâĂŹs prospects. 1

Das Gupta, M. and Mari Bhat, P. (1997). Fertility decline and increased manifestation of sex bias in India. *Population Studies*, 51(3):307–315. 3, 9

Das Gupta, M. and Shuzhuo, L. (1999). Gender bias in China, South Korea and India 1920–1990: Effects of war, famine and fertility decline. *Development and Change*, 30(3):619–652. 8, 9

Dommaraju, P. and Jones, G. (2011). Divorce trends in Asia. *Asian Journal of Social Science*, 39(6):725–750. 18

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. 1, 9

Ebenstein, A. (2011). Estimating a dynamic model of sex selection in China. *Demography*, 48(2):783. 3

Ebenstein, A. (2014). Patrilocality and missing women. *Working Paper*. 3, 8, 9

Ebenstein, A. and Leung, S. (2010). Son preference and access to social insurance: Evidence from China's rural pension program. *Population and Development Review*, 36(1):47–70. 3, 8, 9

Ebrey, P. B. (1993). *The Inner Quarters: Marriage and the Lives of Chinese Women in the Sung Period*. Berkeley: University of California Press. 8

Feeney, G. and Wang, F. (1993). Parity progression and birth intervals in China: The influence of policy in hastening fertility decline. *Population and Development Review*, 19(1):61–101. 1, 2

Freedman, R., Zhenyu, X., Bohua, L., and Lavely, W. (1988). Local area variations in reproductive behavior in the People's Republic of China, 1973–1982. *Population Studies*, 42(1):39–57. 7

Gakidou, E. and King, G. (2006). Death by survey: Estimating adult mortality without selection bias from sibling survival data. *Demography*, 43(3):569–585. 11

Galor, O. and Weil, D. N. (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *American Economic Review*, 90(4):806–828. 1, 36

Glasier, A., Gülmezoglu, A. M., Schmid, G. P., Moreno, C. G., and Van Look, P. F. (2006). Sexual and reproductive health: A matter of life and death. *Lancet*, 368(9547):1595–1607. 2

Goodkind, D. (2011). Child underreporting, fertility, and sex ratio imbalance in China. *Demography*, 48(1):291–316. 12

Goodkind, D. (2017). The astonishing population averted by China's birth restrictions: Estimates, nightmares, and reprogrammed ambitions. *Demography*, 54(4):1375–1400. 2

Greene, M. E. and Merrick, T. (2005). Poverty reduction: Does reproductive health matter? *The World Bank HNP Discussion Paper vol. 33399*. 8, 37

Greenhalgh, S. (2008). *Just One Child: Science and Policy in Deng's China*. Berkeley: University of California Press. 6, 7

Greenhalgh, S. and Li, J. (1993). *Engendering Reproductive Practice in Peasant China: The Political Roots of the Rising Sex Ratios at Birth*. New York: Population Council. 2, 3, 8

Greenhalgh, S. and Winckler, E. A. (2005). *Governing China's Population: From Leninist to Neoliberal Biopolitics*. Stanford: Stanford University Press. 2, 4, 6, 7, 36

Gupta, M. D. (2005). Explaining Asia's "missing women": A new look at the data. *Population and Development Review*, 31(3):529–535. 9

Hazan, M. and Berdugo, B. (2002). Child Labour, Fertility, and Economic Growth. *The Economic Journal*, 112(482):810–828. 1

Howden, D. and Zhou, Y. (2014). C hina's one-child policy: Some unintended consequences. *Economic Affairs*, 34(3):353–369. 1

Hull, T. H. (1990). Recent trends in sex ratios at birth in China. *Population and Development Review*, 16(1):63–83. 9

Institute of Social Science (2014). *China Family Panel Study*. Beijing: Peking University. 12

Jain, A. K. (2011). Measuring the effect of fertility decline on the maternal mortality ratio. *Studies in Family Planning*, 42(4):247–260. 36

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics*, 7(1):63–88. 8

Jayachandran, S. (2017). Fertility decline and missing women. *American Economic Journal: Applied Economics*, 9(1):118–139. 1, 3, 9

Jayachandran, S. and Kuziemko, I. (2011). Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India. *Quarterly Journal of Economics*, 126(3):1485–1538. 9, 26, 34

Jensen, R. T. (2003). Equal treatment, unequal outcomes? Generating sex inequality through fertility behaviour. *Working paper*. 25, 26, 27

Jimmerson, J. (1990). Female infanticide in China: An examination of cultural and legal norms. *Pacific Basin Law Journal*, 8(1). 8

Johansson, S. and Nygren, O. (1991). The missing girls of China: A new demographic account. *The Population and Development Review*, pages 35–51. 4

Joshi, S. and Schultz, T. P. (2013). Family planning and women's and children's health: Long-term consequences of an outreach program in Matlab, Bangladesh. *Demography*, 50(1):149–180. 37

Jowett, A. (1986). China: Population change and population control. *GeoJournal*, 12(4):349–363. 7

Kalemli-Ozcan, S. (2002). Does the mortality decline promote economic growth? *Journal of Economic Growth*, 7(4):411–439. 36

Kane, P. (1987). *The Second Billion: Population and Family Planning in China*. Ringwood, Australia: Penguin. 19

Keynes, J. M. and Volcker, P. A. (1920). *The economic consequences of the peace*, volume 178. Springer. 1

King, M. T. (2014). *Between Birth and Death: Female Infanticide in Nineteenth-Century China*. Stanford: Stanford University Press. 4, 8

Kohler, H.-P. and Philipov, D. (2001). Variance effects in the Bongaarts-Feeney formula. *Demography*, 38(1):1–16. 24

Kremer, M. and Chen, D. L. (2002). Income Distribution Dynamics with Endogenous Fertility. *Journal of Economic Growth*, 7(3):227–258. 1

Langer, W. L. (1974). Infanticide: A historical survey. *Journal of Psychohistory*, 1(3):353. 8

Lavely, W. and Freedman, R. (1990). The origins of the Chinese fertility decline. *Demography*, 27(3):357–367. 2, 10

Lavy, V. and Zablotsky, A. (2011). Mother's schooling and fertility under low female labor force participation: Evidence from a natural experiment. *National Bureau of Economic Research Working Paper*. 36

Lee, B. J. (1981). *Female Infanticide in China*, volume 8. New York: Berghahn Books. 8

Lee, J. and Wang, F. (1999). Malthusian models and Chinese realities: The Chinese demographic system 1700–2000. *Population and Development Review*, 25(1):33–65. 8

Lee, R. (2003). The Demographic Transition: Three Centuries of Fundamental Change. *Journal of Economic Perspectives*, 17(4):167–190. 1

Li, N., Feldman, M. W., and Tuljapurkar, S. (2000). Sex ratio at birth and son preference. *Mathematical Population Studies*, 8(1):91–107. 3, 26

Lu, W. and Mungello, D. (2010). Drowning girls in China: Female infanticide since 1650. *Journal of Asian Studies*, 69(1):230. 4, 8

Menken, J. and Rahman, M. O. (2001). Reproductive health. *International Public Health: Diseases, Programs, Systems, and Policies. Gaithersburg, MD: Aspen Publishers, Inc*, pages 79–138. 36

Merli, M. G. and Raftery, A. E. (2000). Are births underreported in rural China? manipulation of statistical records in response to China's population policies. *Demography*, 37(1):109–126. 12

Miller, G. (2010). Contraception as development? New evidence from family planning in Colombia. *Economic Journal*, 120(545):709–736. 37

Miller, G. and Babiarz, K. S. (2016). Family planning: Program effects. *Population and Development Review*, 42(1):7–26. 1, 3, 5

Milligan, K. (2005). Subsidizing the Stork: New Evidence on Tax Incentives and Fertility. *The Review of Economics and Statistics*, 87(3):539–555. 1

Moav, O. (2005). Cheap Children and the Persistence of Poverty. *The Economic Journal*, 115(500):88–110. 1

Mosher, S. W. (2008). *Broken Earth*. New York: Simon and Schuster. 2

Mungello, D. E. (2008). *Drowning Girls in China: Female Infanticide in China Since 1650*. Lanham, MD: Rowman & Littlefield Publishers. 8

Park, C. B. and Cho, N.-H. (1995). Consequences of son preference in a low-fertility society: Imbalance of the sex ratio at birth in Korea. *Population and Development Review*, pages 59–84. 26

Pop-Eleches, C. (2006). The impact of an abortion ban on socioeconomic outcomes of children: Evidence from Romania. *Journal of Political Economy*, 114(4):744–773. 37

Preston, S. H. et al. (1978). *The Effects of Infant and Child Mortality on Fertility*. New York: Academic Press. 2, 15

Pritchett, L. H. (1994). Desired fertility and the impact of population policies. *Population and Development Review*, 20(1):1–55. 1, 5, 26, 36

Raute, A. (2019). Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits. *Journal of Public Economics*, 169:203–222. 1

Robinson, W. C. and Ross, J. A. (2007). *The Global Family Planning Revolution: Three Decades of Population Policies and Programs*. Washington DC: The World Bank. 35

Rosenzweig, M. R. and Schultz, T. P. (1983). Consumer demand and household production: The relationship between fertility and child mortality. *American Economic Review*, 73(2):38–42. 2, 15

Scharping, T. (2013). *Birth Control in China 1949-2000: Population policy and demographic development*. Routledge. 2, 6, 22

Schultz, T. P. (1985). Changing world prices, women's wages, and the fertility transition: Sweden, 1860-1910. *Journal of Political Economy*, 93(6):1126–1154. 15, 36

Schultz, T. P. (1994). Human capital, family planning, and their effects on population growth. *The American Economic Review*, 84(2):255–260. 1

Schultz, T. W. (1973). The Value of Children: An Economic Perspective. *Journal of Political Economy*, 81(2):S2–S13. 1

Song, J., Tuan, C.-H., and Yu, J. (1985). *Population Control in China: Theory and Applications.* New York: Praeger. 6

Song, S. (2012). Does famine influence sex ratio at birth? evidence from the 1959–1961 Great Leap Forward famine in China. *Proceedings of the Royal Society of London B: Biological Sciences*. 8

United Nation Population Division (2017). *World Fertility Data*. New York: United Nations, Department of Economic and Social Affairs, Population Division. 45

Van Hook, J. and Altman, C. E. (2013). Using discrete-time event history fertility models to simulate total fertility rates and other fertility measures. *Population Research and Policy Review*, 32(4):585–610. 5, 20, 22, 23, 57

Wang, F. (2012). Family planning policy in china: Measurement and impact on fertility. 6

Wang, F., Gu, B., and Cai, Y. (2016). The end of China's One-Child Policy. *Studies in Family Planning*, 47(1):83–86. 5

White, T. (2006). *China's Longest Campaign: Birth Planning in the People's Republic, 1949-2005*. Ithaca, NY: Cornell University Press. 2, 3, 6, 7, 8, 10, 15, 36

Whyte, M. K., Wang, F., and Cai, Y. (2015). Challenging myths about China's One Child Policy. *The China Journal*, (74):144–159. 7, 8

Wilmoth, J. R., Andreev, K., Jdanov, D., Glei, D. A., Boe, C., Bubenheim, M., Philipov, D., Shkolnikov, V., and Vachon, P. (2007). Methods protocol for the human mortality database. *University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock*, 9:10–11. 24

Winikoff, B. and Sullivan, M. (1987). Assessing the role of family planning in reducing maternal mortality. *Studies in Family Planning*, 18(3):128–143. 36

Wolf, A. P. and Huang, C.-s. (1980). Marriage and adoption in China 1845-1945. 4, 8

Yamaguchi, K. (1989). A formal theory for male-preferring stopping rules of childbearing: Sex differences in birth order and in the number of siblings. *Demography*, 26(3):451–465. 4, 25, 26, 27

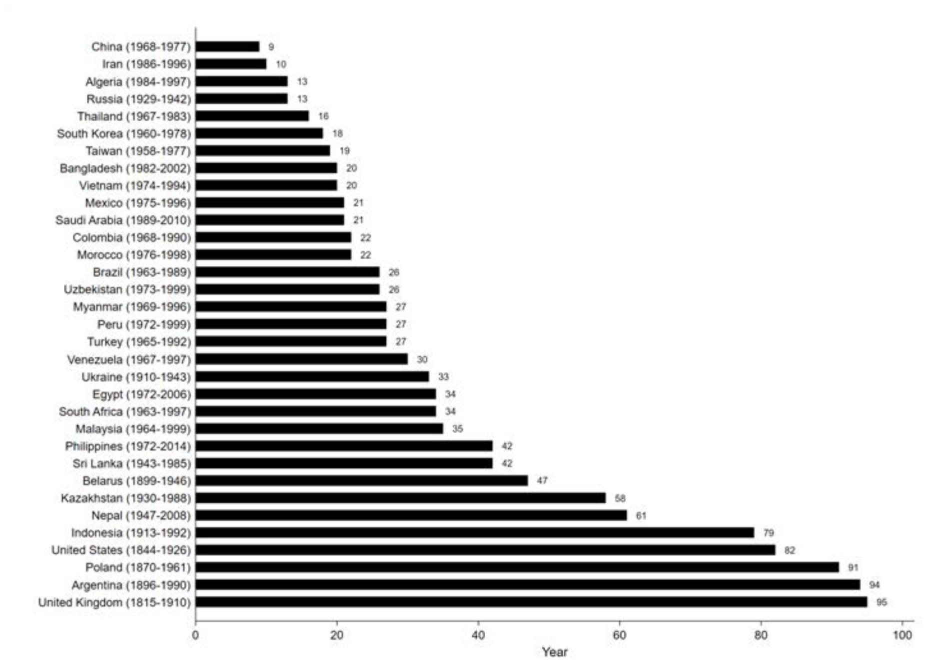
Yi, Z. (1996). Is fertility in China in 1991–92 far below replacement level? *Population Studies*, 50(1):27–34. 12

Yi, Z., Ping, T., Baochang, G., Yi, X., Bohua, L., and Yongpiing, L. (1993). Causes and implications of the recent increase in the reported sex ratio at birth in China. *Population and Development Review*, 19(2):283–302. 9

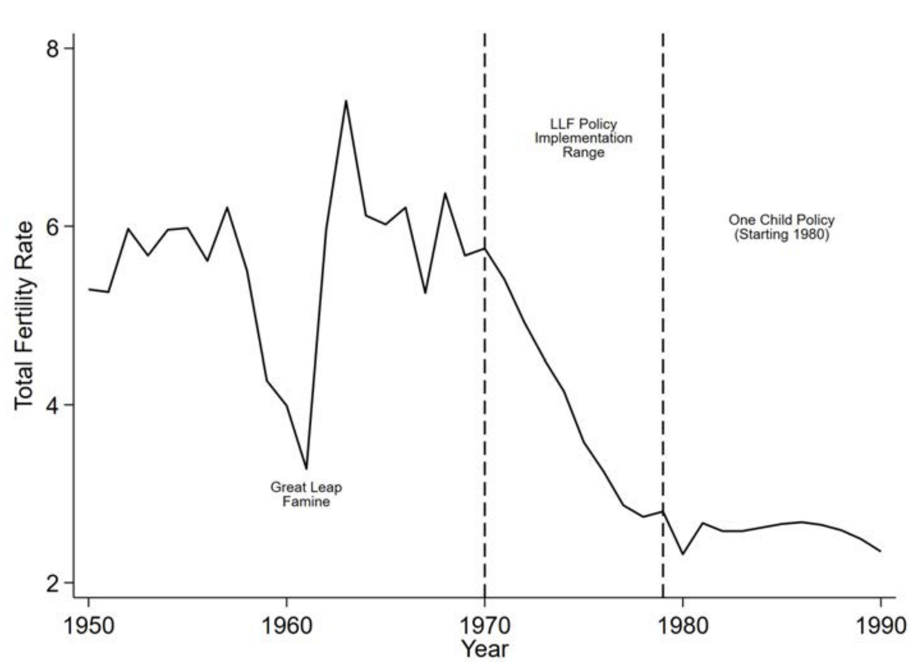
Zhang, G. and Zhao, Z. (2006). Reexamining China's fertility puzzle: Data collection and quality over the last two decades. *Population and Development Review*, 32(2):293–321. 12

Zheng, L., Jian, S., Liu, Z., and Song, J. (1981). *China's Population: Problems and Prospects*. Beijing: New World Press. 6

Figure 1.
Total Fertility Rate: China 1950-1990 and Historical Fertility Transitions



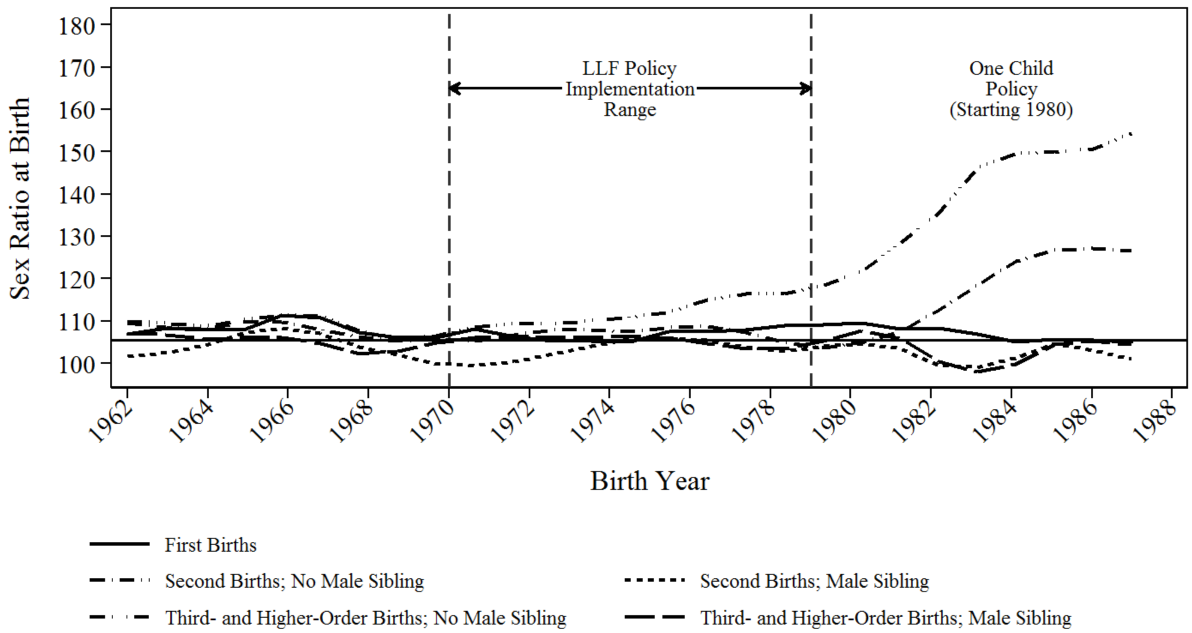
Panel A



Panel B

Panel A shows the number of years for countries to transition from an above 6 total fertility rate (TFR) to a below 3 TFR. Sample constructed from Gapminder v12 TFR and v6 Population data and is restricted to countries with at least 0.25% of the world population in the initial transition year. For each country, the start year is the last year in which TFR $\geq$ 6 and the end year is the first year in which TFR $\leq$ 3. Panel B shows the Total Fertility Rate of China from 1950-1990 United Nation Population Division (2017).

Figure 2.
Sex Ratio at Birth by Parity and Sibship Sex Composition: China, 1962-1987

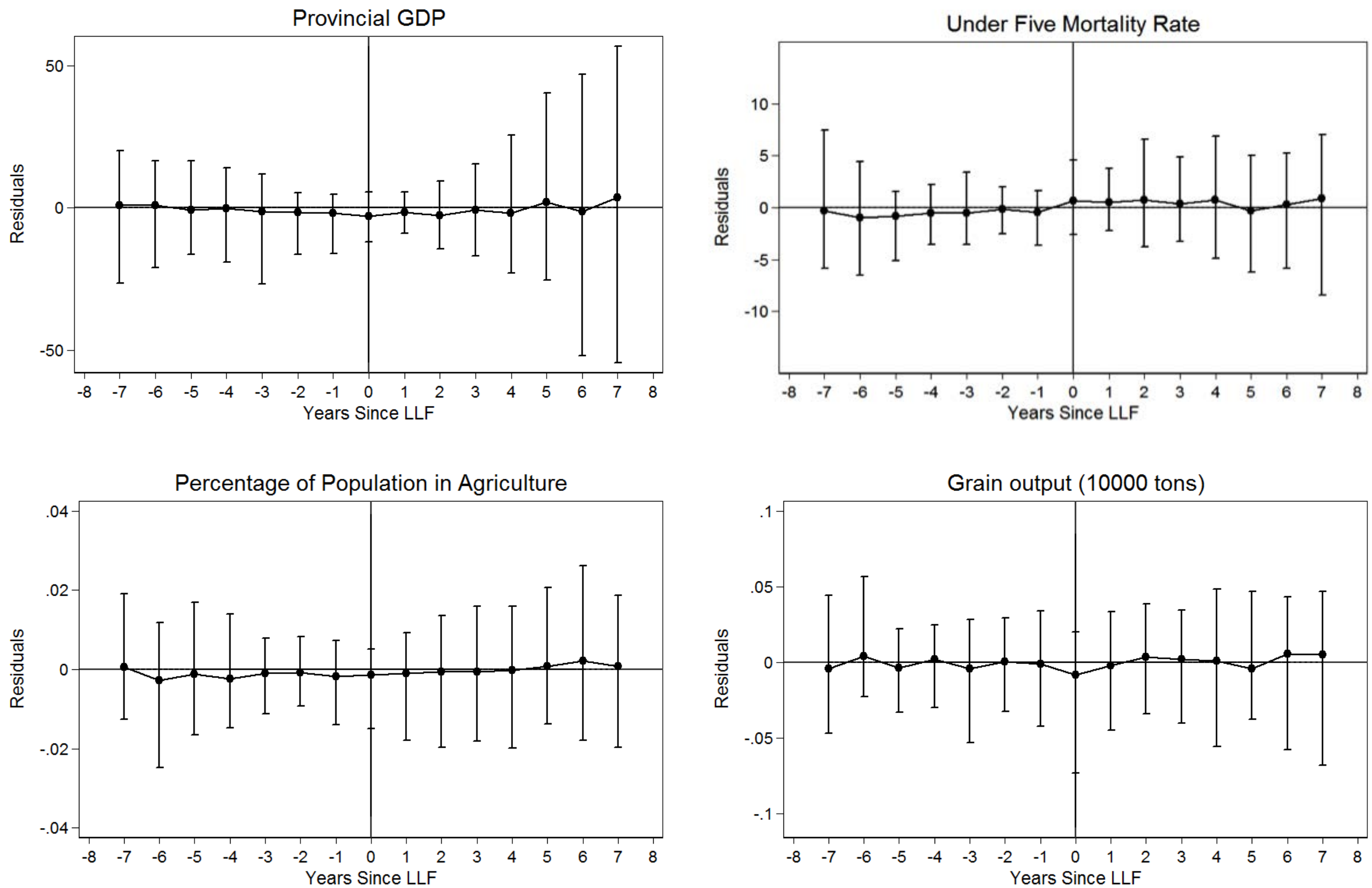


Note: Figure shows sex ratios at birth by parity and sex composition of previous births (parents with and without a previously born boy). Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception

Figure 3.
Years of Later, Longer, Fewer, Campaign Implementation

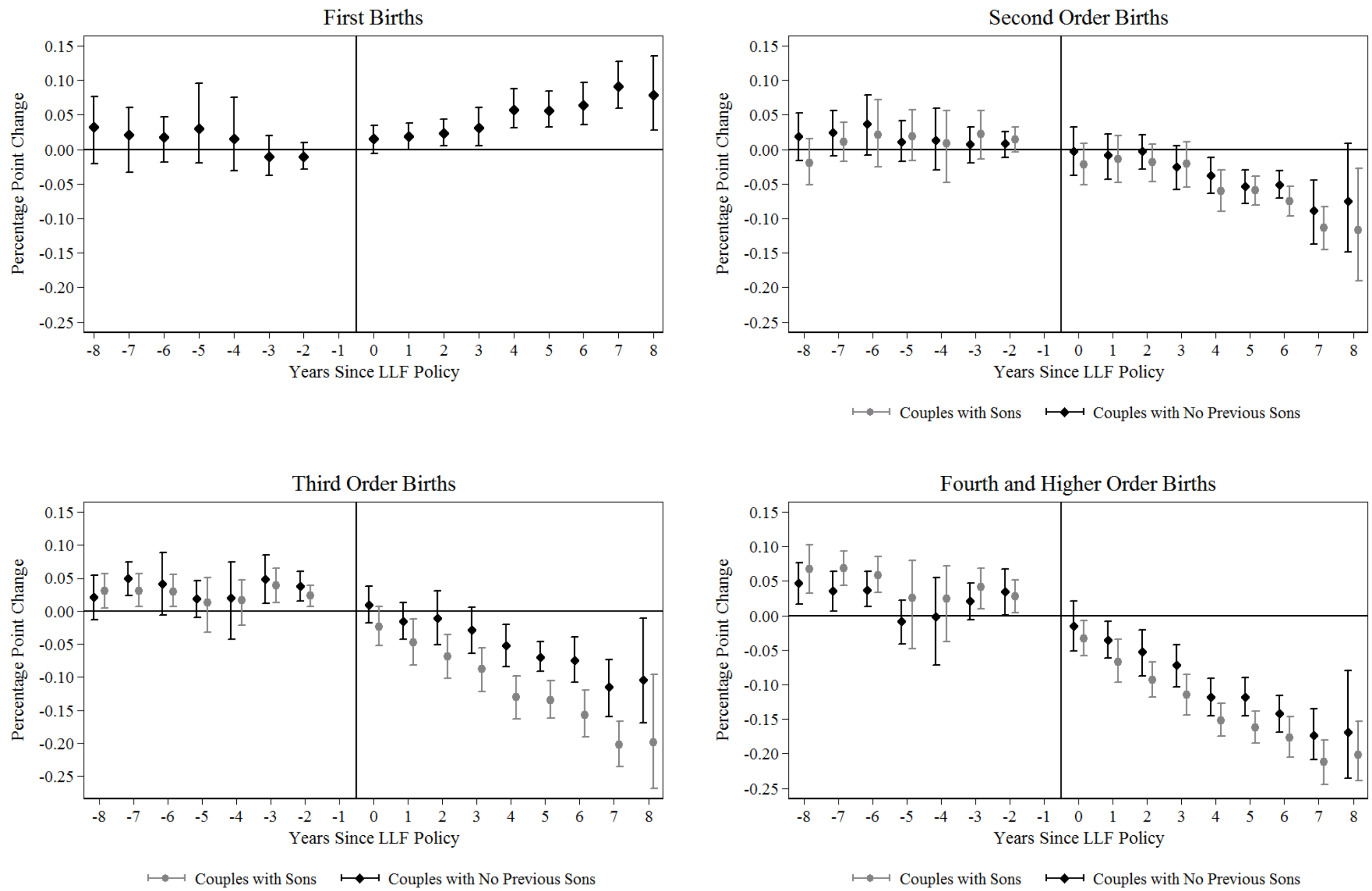
Note: Figure shows the year in which province-level committees were formed to implement national birth-planing policies. Data: Digitized records from provincial public health archives (*Weishengzhi*) and historical provincial annals.

Figure 4.
LLF Timing and Determinants of Fertility Demand



Note: Figure plots mean and 95% range of residuals after conditioning established determinants of the demand for children on province and calendar year fixed effects by event year (normalizing the year of LLF implementation in each province to be event year zero). Data: Digitized provincial public health archive records (*Weishengzhi*), and National Bureau of Statistics of China.

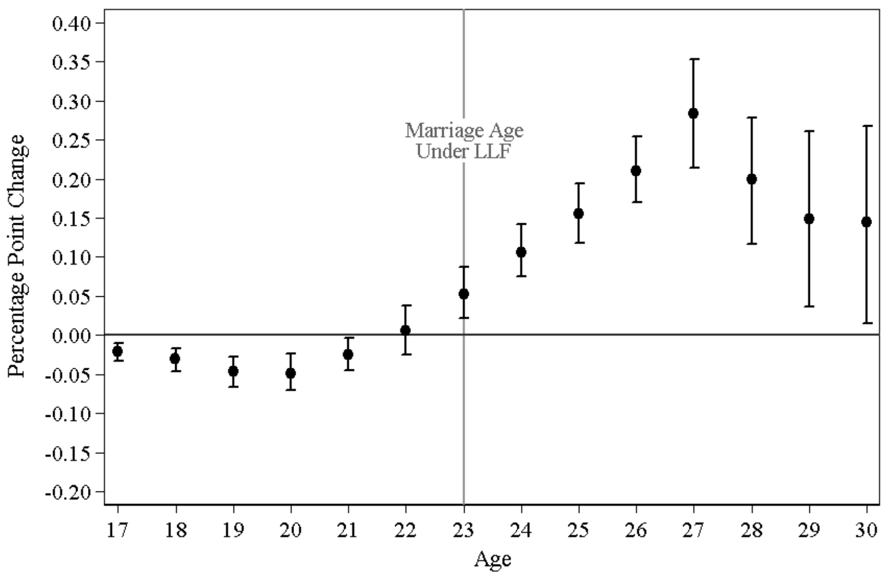
Figure 5.
The LLF Policy and the Annual Risk of Parity Progression by Parity and Sibship Sex Composition



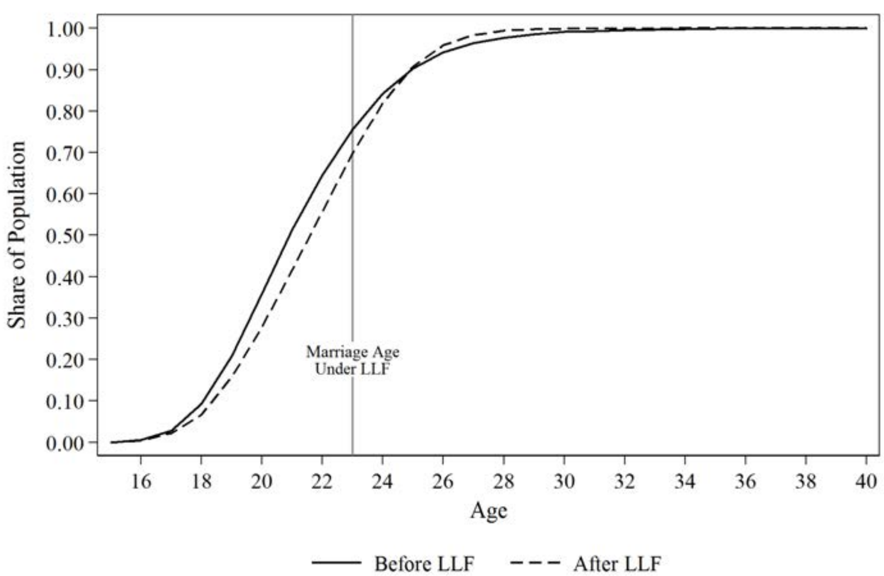
Note: Figure shows estimates from Equation 1, which estimates the change in the probability of a parity-specific birth among married women by sex composition of previous births and event year. We condition on maternal and household characteristics (a mother's highest level of education, her age at marriage, and the ethnicity of her household head) as well as province-year characteristics (five-year average under-5 mortality rate, provincial GDP, gross agricultural output, grain production, and the proportion of the population classified as rural), provincial fixed effects, and calendar year fixed effects. Confidence intervals estimated using the wild bootstrap method with 1000 replications (Cameron et al., 2008). Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.

Figure 6.
The LLF Policy and Age at Marriage

(a) Panel A: Marginal Effect of the LLF Policy on the Age-Specific Probability of Marriage



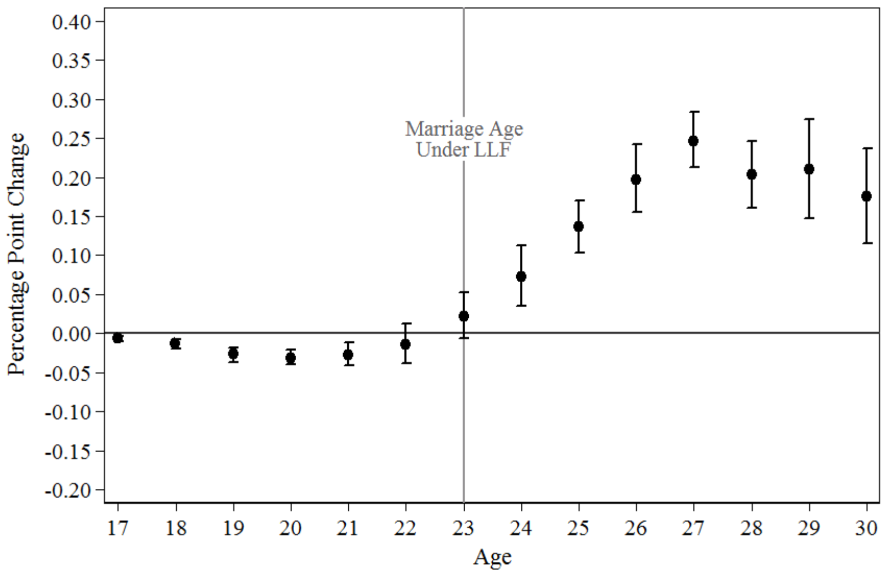
(b) Panel B: Implied Proportion of Population Married with and without LLF



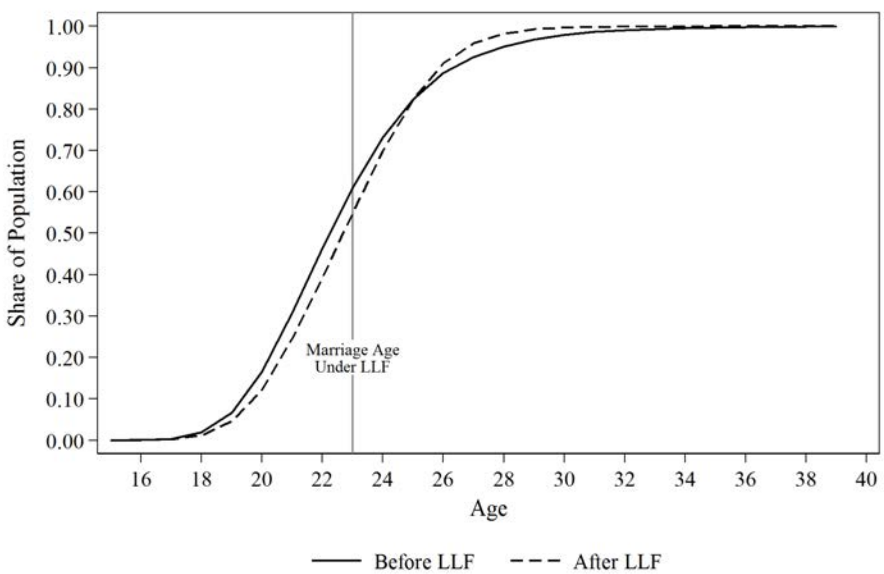
Note: Figure 6, Panel A shows discrete-time hazard model estimates from Equation 2 for age of marriage among unmarried women aged 15 and older. Estimates and 95% confidence intervals are linear combinations of indicators for the LLF policy and the interactions between the policy and an individual's age. We condition on maternal and household characteristics (a mother's highest level of education, her age at marriage, and the ethnicity of her household head) as well as province-year characteristics (five-year average under-5 mortality rate, provincial GDP, gross agricultural output, grain production, and the proportion of the population classified as rural), provincial fixed effects, and calendar year fixed effects. Panel B shows the implied cumulative proportion of women married by single year of age. Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.

Figure 7.
The LLF Policy and Age at First Birth

(a) Panel A: Marginal Effect of the LLF Policy on the Age-Specific Probability First Birth

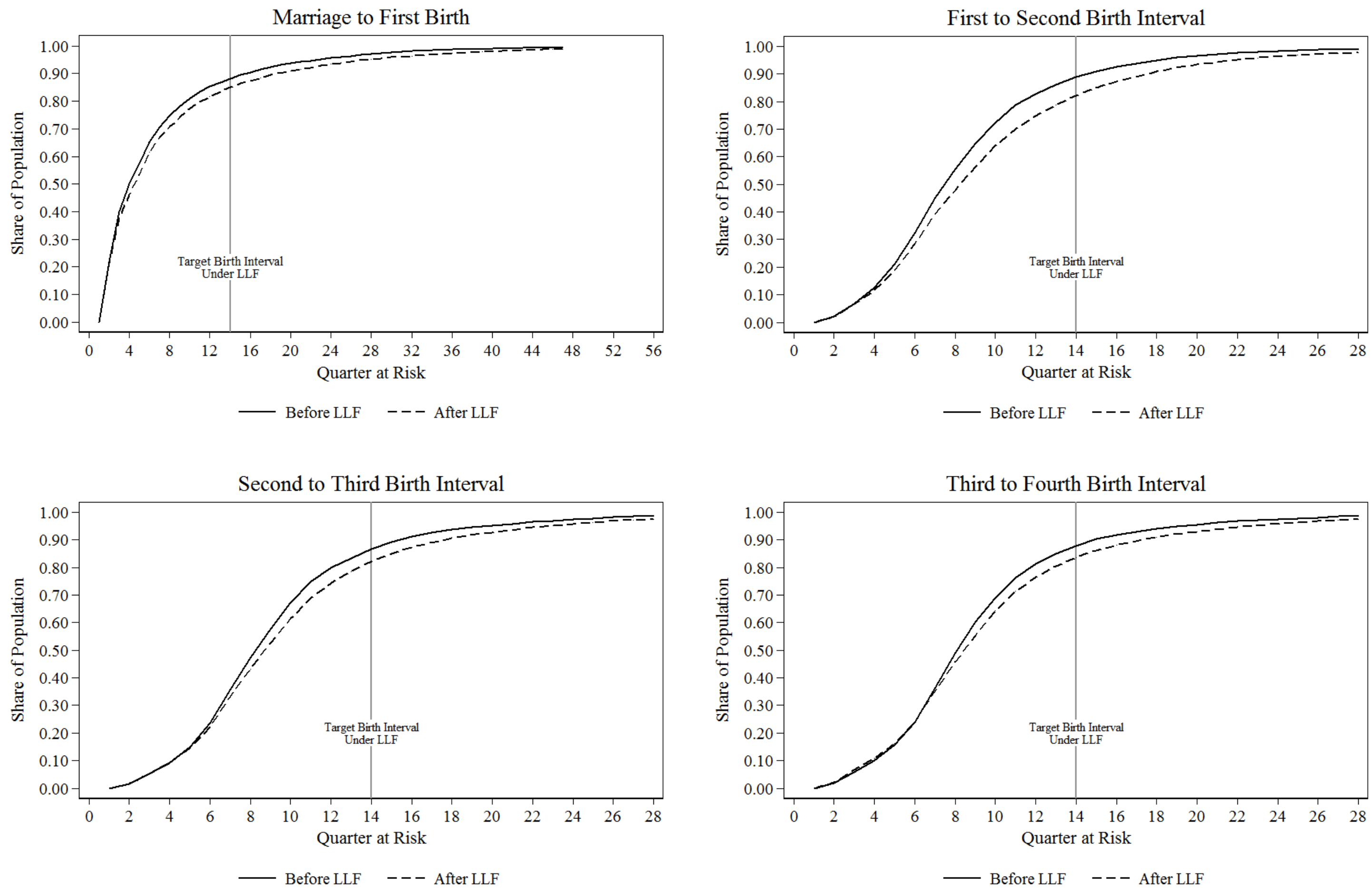


(b) Panel B: Implied Proportion of Women Having First Birth with and without LLF



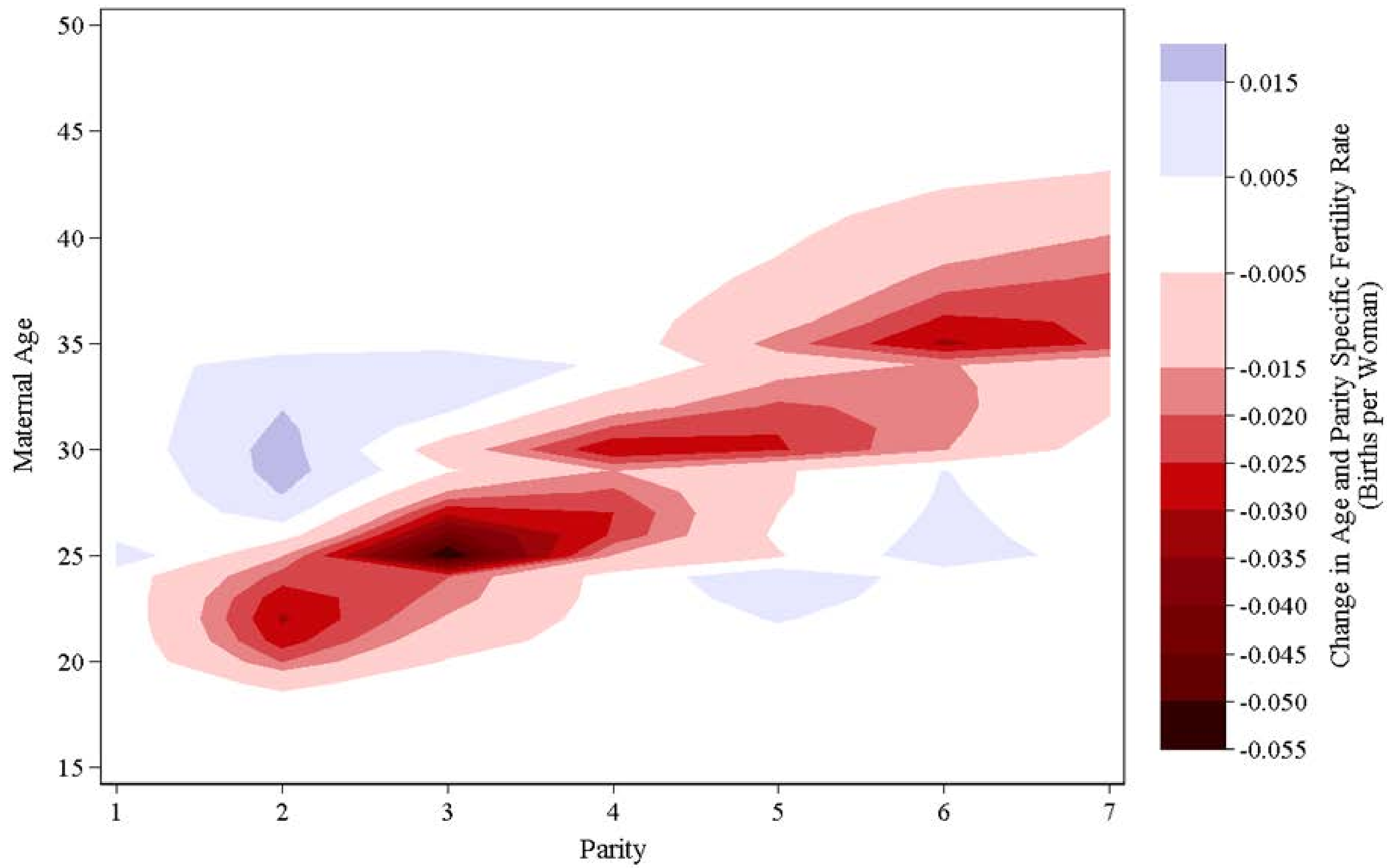
Note: Figure 7, Panel A shows the results of a discrete-time hazard model estimates from Equation 2 for age of first parity birth among women aged 15 and older. Estimates and 95% confidence intervals are linear combinations of indicators for the LLF policy and the interactions between the policy and an individual's age. We condition on maternal and household characteristics (a mother's highest level of education, her age at marriage, and the ethnicity of her household head) as well as province-year characteristics (five-year average under-5 mortality rate, provincial GDP, gross agricultural output, grain production, and the proportion of the population classified as rural), provincial fixed effects, and calendar year fixed effects. Panel B shows the implied cumulative proportion of women having had a first birth by single year of age. Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.

Figure 8.
The LLF Policy and Birth Interval Length



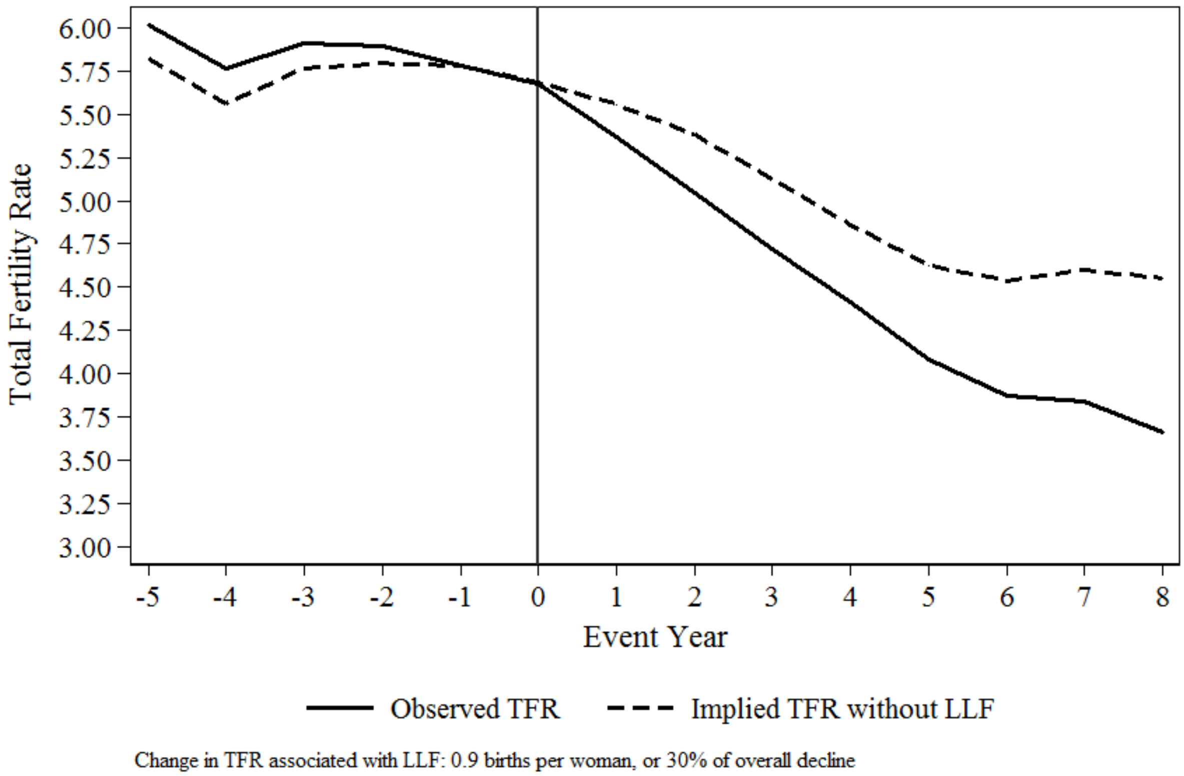
Note: Figure 8 shows the implied cumulative proportion of the population of parity n mothers advancing to parity $n+1$ birth by quarter at risk. Using a woman-quarter sample of mothers in which each mother enters the risk set 3 quarters after a parity n birth and exits the risk set in the period in which a parity $n+1$ occurs, we calculate the probability of parity progression in each quarter since the previous birth with and without LLF implied by estimates from Equation 3. Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.

Figure 9.
Changes in Age- and Parity-Specific Fertility Rates During the LLF Period



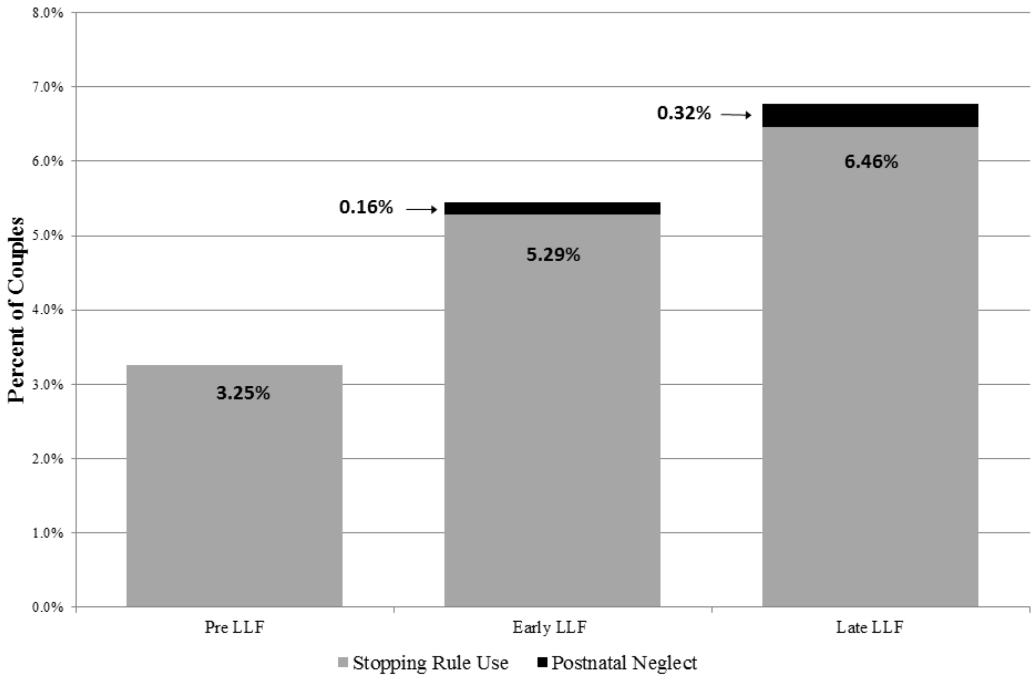
Note: Figure 9 shows the implied change in age- and parity-specific fertility rates during the LLF period. Using age- and parity-specific fertility rates implied by 4, we subtract fertility rates for the year of LLF implementation from equivalent rates in event year 8 to show the age and parity specific pattern of LLF's effects on fertility. Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.

Figure 10.
The LLF Policy and Total Fertility Rate by Event Year



Note: Figure 10 shows the observed Total Fertility Rate and counterfactual TFR, which is calculated by subtracting the policy-driven change in TFR implied by estimates of Equation 4 from observed TFR in each event year. Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.

Figure 11.
Summary of the LLF Policy and Prevalence of Sex Composition Strategies by Type



Note: Figure shows the share of couples using male-biased fertility stopping rules and postnatal sex selection as fertility strategies favoring sons by period, implied by estimates from Equations 9 and 10. Specifically, we recover the prevalence of stopping rule use due to the LLF policy by subtracting our estimates of postnatal selection (Equation 10) from the share of couples using any sex composition strategy (Equation 9) in each parity and sibship sex composition group. We then weight by the relative proportion of couples in each parity and sex composition group (in each period) Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.

Table 1.
Summary Statistics

	Overall	Pre-LLF Event Years 0 and Prior	Early LLF Event Years 1-4	Late LLF Event Years 5+
Birth Characteristics				
Percent Male (%)	51.63%	51.57%	51.72%	51.66%
Marriage to First Birth Interval (Months)	21.4	24.9	22.8	17.7
First Birth to Second Birth Interval (Months)	29.1	30.5	29.4	27.4
Second Birth to Third Birth Interval (Months)	31.4	32.7	32.4	29.3
Third Birth to Fourth Birth Interval (Months)	32.3	33.2	32.4	30.2
Maternal Characteristics				
Mean Age at Marriage (Years)	19.57	19.01	19.74	20.64
Mean Maternal Age at First Birth (Years)	22.49	21.83	22.67	23.54
Share Illiterate	48.8%	51.8%	46.8%	44.4%
Share Semi Literate	11.4%	11.1%	11.5%	11.9%
Share with Primary School Education	29.8%	28.3%	31.5%	31.2%
Share with Middle School Education or Above	10%	8.9%	10.2%	12.5%
Provincial Characteristics				
Under 5 Mortality Rate (per 1000)	20.27	25.49	16.27	13.30
Grain output (1000 tons)	3.02	2.86	3.14	3.25
Agricultural Production (1000 tons)	14.4	8.55	16.2	25.9
Provincial GDP (100M Yuan)	81.3	61.86	87.29	119.35
Share of Population Rural	79.5%	78.9%	79.8%	80.7%
Sample Size				
Mother-Year Observations	1,279,362	505,485	404,335	369,542
All Births	292,756	146,534	86,752	61,470
Parity 1 Births	80,998	37,736	22,926	20,336
Parity 2 Births	73,611	35,060	21,780	16,771
Parity 3 Births	57,600	29,076	16,861	11,663
Parity 4+ Births	82,547	44,662	25,185	12,700

Note: Summary statistics calculated among mother-years and births used in our analysis. Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, and National Bureau of Statistics of China.

Table 2.
Implied Age-Specific Fertility Rates, Total Fertility Rates, and Tempo Adjusted Fertility Rates by Event Year

	Age Specific Fertility Rates					Total Fertility Rate	
	15-19	20-24	25-29	30-34	35+	TFR	TFR'
5 Years Prior To LLF	0.20	1.62	1.87	1.34	1.10	6.14	6.14
4 Years Prior To LLF	0.19	1.59	1.88	1.40	1.08	6.14	6.14
3 Years Prior To LLF	0.19	1.53	1.89	1.38	1.10	6.09	6.09
2 Years Prior To LLF	0.20	1.56	1.85	1.36	1.08	6.04	6.04
1 Year Prior To LLF	0.19	1.50	1.87	1.36	1.02	5.94	5.94
Year of LLF	0.21	1.51	1.91	1.30	1.01	5.93	5.93
1 Year After LLF	0.22	1.50	1.84	1.27	0.93	5.76	5.81
2 Years After LLF	0.22	1.46	1.80	1.26	0.87	5.60	5.66
3 Years After LLF	0.23	1.48	1.79	1.20	0.84	5.54	5.60
4 Years After LLF	0.22	1.47	1.77	1.21	0.82	5.49	5.55
5 Years After LLF	0.21	1.39	1.80	1.24	0.76	5.40	5.46
6 Years After LLF	0.21	1.38	1.74	1.19	0.74	5.27	5.33
7 Years After LLF	0.22	1.38	1.66	1.19	0.74	5.18	5.24
8 Years After LLF	0.19	1.34	1.71	1.13	0.68	5.05	5.11

Note: Table shows regression adjusted age-specific fertility rates for each event year implied by Equation 4 (Columns 1-5). Following the general method developed in Van Hook and Altman (2013), we use regression estimates to predict birth rates by maternal age and parity for each event year, holding maternal and household characteristics (a mother's highest level of education, her age at marriage, and the ethnicity of her household head), province-year characteristics (five-year average under-5 mortality rate, provincial GDP, gross agricultural output, grain production, and the proportion of the population classified as rural), calendar year and province fix effects constant at reference year values (event year -1) to isolate the effect of the LLF policy. Summing across age groups, Column 6 shows the implied Total Fertility Rate. We then adjust the TFR decline for changes in the age at childbearing, following Bongaarts and Feeney (1998), to estimate the change in quantum fertility (see Appendix Table A6). Column 7 shows these tempo-adjusted TFRs. Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.

Table 3.
The LLF Policy and Sex-Based Discontinuation of Childbearing

	Couples with No Previous Sons			Couples with At Least One Previous Son		
	Pre-LLF	Early LLF	Late LLF	Pre-LLF	Early LLF	Late LLF
First Parity Births	0.005***	0.024***	0.088***	-	-	-
	[0.002 - 0.009]	[0.014 - 0.034]	[0.051 - 0.129]	-	-	-
Second Parity Births	0.080***	0.206***	0.257***	0.013**	0.030**	0.053***
	[0.055 - 0.107]	[0.150 - 0.263]	[0.202 - 0.306]	[0.002 - 0.023]	[0.008 - 0.052]	[0.019 - 0.092]
Third + Parity Births	0.136***	0.270***	0.310***	0.052***	0.064***	0.072***
	[0.103 - 0.170]	[0.202 - 0.342]	[0.238 - 0.374]	[0.043 - 0.061]	[0.051 - 0.077]	[0.045 - 0.104]

Note: Each row shows the parity-specific marginal effect of having a male birth in each LLF period on the likelihood of discontinuing childbearing (compared to mothers of the same parity in the same period having a female birth). Ordinary least squares regressions shown in Equation 9 are stratified by parity and control for maternal and household characteristics (a mother's highest level of education, her age at marriage, and the ethnicity of her household head) as well as province-year characteristics (five-year average under-5 mortality rate, provincial GDP, gross agricultural output, grain production, and the proportion of the population classified as rural), calendar year fixed effects and provincial fixed effects. Confidence sets estimated using the wild bootstrap method with 1000 replications (Cameron et al., 2008). Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.

Table 4.
The LLF Policy and the Probability of a Male Birth
Among Couples With No Previous Sons

	Pre-LLF	Early LLF	Late LLF
First Parity Births	-	-	-
	-	-	-
Second Parity Births	0.006	0.015**	-0.002
	[-0.008 - 0.020]	[0.001 - 0.028]	[-0.034 - 0.027]
Third + Parity Births	0.002	0.006	0.023**
	[-0.008 - 0.011]	[-0.008 - 0.021]	[0.004 - 0.043]

Note: Each row shows the parity-specific marginal effect of a couple not having any previously born sons on the likelihood a particular birth is male in each LLF period (compared to otherwise similar births occurring to couples with a previously born son). In other words, the coefficients show the increases over event time in the effect of sibship sex composition on the probability of a male birth. Ordinary least squares regressions described in Equation 10 are stratified by parity, and control for maternal characteristics, province-year characteristics, calendar year fixed effects and provincial fixed effects. 95% confidence sets estimated using the wild bootstrap method with 1000 replications (Cameron et al., 2008). Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.

Table 5.
The LLF Policy and Male-Biased Stopping Rule Prevalence

	Couples with No Previous Sons			Couples with At Least One Previous Son		
	Pre-LLF	Early LLF	Late LLF	Pre-LLF	Early LLF	Late LLF
First Parity Births	0.003	0.012	0.046	-	-	-
Second Parity Births	0.041	0.091	0.132	0.007	0.016	0.028
Third + Parity Births	0.070	0.139	0.137	0.027	0.033	0.037

Note: Each cell shows the prevalence of male-biased fertility stopping rule use by parity and sibship sex composition implied by estimates of excess discontinuation in Equation 9 and by estimates of postnatal selection in Equation 10. These are calculated by subtracting estimates of postnatal selection for each parity- and sibship sex composition group from estimates of the prevalence of excess fertility discontinuation among couples in each group. Data: 1988 "Two-Per-Thousand" National Survey of Fertility and Contraception, digitized provincial public health archive records, National Bureau of Statistics of China, and the China Family Panel Survey.