The Impact of Margaret Sanger's Birth Control Clinics on Early 20th Century U.S. Fertility and Mortality†

Stefan Bauernschuster[a,b,c], Michael Grimm[a,c,d], Cathy M. Hajo[e]

[a] *University of Passau, Germany*
[b] *CESifo, Munich*
[c] *IZA, Institute for the Study of Labor, Bonn*
[d] *RWI-Research Network, Essen*
[e] *Ramapo College, New Jersey, USA*

April 1, 2025

Abstract

In 1916, Margaret Sanger established the first birth control clinic in U.S. history. From the mid-1920s, "Sanger clinics" spread over the entire U.S. These clinics advised mothers on contraception, mainly by fitting diaphragms and explicitly instructing women how to use them effectively. Combining newly digitized data on the roll-out of these clinics, full-count Census data, and historical vital statistics, we find that the clinics accounted for 5.4–6.1 percent of the overall fertility decline across the U.S. until 1940. By increasing the spacing between births and reducing fertility, the clinics generated important health effects and reduced stillbirths and infant mortality.

JEL Classification: D10, J13, J23, N32, O12
Keywords: birth control, fertility, mortality, Margaret Sanger, demographic transition

†Stefan Bauernschuster (corresponding author), School of Business Administration, Economics, and Information Systems, University of Passau, Innstrasse 27, 94032 Passau, Germany, email: stefan.bauernschuster@uni-passau.de. Michael Grimm, School of Business Administration, Economics, and Information Systems, University of Passau, Innstrasse 27, 94032 Passau, Germany, email: michael.grimm@uni-passau.de. Cathy M. Hajo, Ramapo College of New Jersey, 505 Ramapo Valley Road, Mahwah, NJ 07430, USA, chajo@ramapo.edu. We are thankful to Heather Royer, Stephen L. Ross, and three anonymous referees for very constructive feedback and suggestions. For helpful comments and exchange, we further would like to thank Philipp Ager, Martha Bailey, Bastien Chabe-Ferret, Shari Eli, Stephan Heblich and Helmut Rainer as well as participants at the 3rd Berlin Workshop on Empirical Public Economics: Gender Economics, the CESifo Area Conference on Labor Economics 2023, the 2023 EUDN conference in Geneva, the ESPE conference in Belgrade, the EALE conference in Bergen, the GDE Conference 2022 in Stuttgart-Hohenheim, the GlaD-DENeB Seminar in Berlin, the 17th North American Meeting of the Urban Economics Association in Toronto as well as the Berlin Applied Micro Seminar, and seminars participants at the University of Connecticut, the University of Duisburg-Essen, the University of Eichstaett-Ingolstadt, the University of Groningen, the University of Hannover, the University of Passau and the University of Wageningen. Julian Hagemann, Edwin Mejía, Angelina Odintsova, Natalia Oviedo, Thomas Schiller, Seraphin Schwarzbauer and Padmaja Kadambi Seshadri provided valuable research assistance.

> *"Margaret Sanger (1879-1966) led a successful campaign from 1914 to 1937 to remove the stigma of obscenity from contraception and to establish a nationwide system of clinics where women could obtain reliable birth control services. She organized research, recruited manufacturers for birth control devices, and won court battles that modified the Comstock laws and laid the groundwork for the formal acceptance of birth control by organized medicine in 1937. After World War II she played key roles in the rise of an international planned parenthood movement and in the development of the birth control pill. Through these achievements she had a greater impact on the world than any other American woman."*
>
> – Reed (1978, p. 67)

1. Introduction

In 1873, the U.S. Congress passed the Comstock Act, making it illegal to mail, ship, or import "obscene" articles, drugs, medicines, and printed materials, including any article or information related to the prevention of conception. As a consequence, contraception was banned under federal and most state statutes (Tone, 2002). Stringent Comstock laws made it impossible in particular for medical professionals to provide contraceptives or offer respective advice without risking their licences and livelihoods (Bailey and Hershbein, 2018). According to Guttmacher (1947), 90 percent of all physicians earning a medical degree before 1920 had not even received any training on contraception at medical school leading women to rely on ill-informed sources, often quacks and charlatans, for ineffective and sometimes even dangerous contraceptives.

In response to these restrictive measures, a counter-movement providing professional and safe birth control for those in need emerged at the beginning of the 20th century. Margaret Sanger was one of the leading figures of this movement. On October 16, 1916, she established the United States' first birth control clinic in New York, Brownsville, Brooklyn—despite facing significant opposition (Engelman, 2011). According to Sanger's idea, these birth control clinics should give women advice on contraception, sexual health and hygiene. Most importantly, by fitting diaphragms and explicitly instructing women how to use these devices effectively, the clinics should offer services not available via any other institution at that time (Hajo, 2010). Providing these services, the birth control activists aimed at putting women in control of contraceptive decisions and thus empowering them to lengthen the interpregnancy interval, allowing them to get closer to the desired family size, thereby improving mothers' and children's health and alleviating poverty in (too) large families. Although the Brownsville clinic was shut down by the police ten days after its opening, similar 'Sanger clinics' began to appear throughout the U.S. from 1923 onward, with over 600 established by 1940 (Hajo, 2010). Despite the historical relevance of Margaret Sanger's early birth control movement, the impact of her birth control clinics has not yet been empirically assessed. Filling this gap should help to enhance our understanding of the historical demographic transition and its broader socio-economic effects in the U.S.

In this paper, we zoom in on the beginnings of the U.S. birth control movement and estimate the impact of the early birth control clinics on fertility, stillbirths, infant mortality, and puerperal deaths. To this end, we have digitized detailed data on the universe of birth control clinics established from 1916 to 1940. We exploit the staggered roll-out of the clinics across U.S. counties and combine the unique clinic data with full-count Census data of 1920, 1930 and 1940, yearly historical vital statistics at the county level, and yearly causes of death data at the city level. The Census data allow us to investigate the effects of birth control clinics on the number of a woman's own children under the age of five living in a household. Using the county level vital statistics, we re-assess the effect of birth control clinics on fertility and estimate the impact on stillbirths and infant mortality. Causes of death data at the city level allow us to inspect the impact of the clinics on puerperal deaths.

We expect that Sanger's birth control clinics empowered women to control fertility, reduced the cost of increasing interpregnancy intervals and thus satisfied a demand for effective contraception. By increasing the interpregnancy intervals, the clinics might not just have reduced fertility but also fetal and infant mortality. This link between birth spacing and infant health, which has been persuasively documented in both historical and recent data (Conde-Agudelo *et al.*, 2006; Knodel and Hermalin, 1984; Molitoris, 2017), can be explained medically but also economically if lower fertility alleviates budget constraints for households and thus increases health investments per child. Additionally, by increasing birth spacing and decreasing fertility rates, birth control clinics might have lowered maternal deaths, especially puerperal deaths.

In the first part of the empirical analysis, we use full-count Census data and exploit differences in the length of exposure to birth control clinics conditional on county and year fixed effects as well as a rich set of socio-economic controls. Our findings reveal statistically significant and economically meaningful effects of birth control clinics on fertility, measured by the number of a woman's own children below the age of five in the household. Specifically, evaluated at the average exposure time of a women at the end of our observation period of 5.77 years, the estimate implies a fertility reduction of 3.7 percent. From 1920 to 1940, birth control clinics thus account for 5 percent of the overall fertility decline across the U.S. Event-study analyses provide evidence for the validity of the key identifying assumption. The findings are robust to alternative specifications in which we control for state-specific cohort fixed effects, add interactions of Census year and variables capturing baseline differences in fertility and the local intensity of the Great Depression, or restrict the sample to big city counties.

In the second part of the empirical analysis, we use yearly county-level vital statistics and exploit the staggered roll-out of birth control clinics across counties over time. To avoid pitfalls of standard two-way fixed effects models in case of heterogeneous treatment effects, we use the interaction weighted estimator proposed by Sun and Abraham (2021). The estimates confirm the negative impact of birth control clinics on fertility. Birth control clinics decreased the local crude birth rate by 3.4 percent within ten years after establishment. Overall, birth control clinics explain 6.1 percent of the total decline of the crude birth rate in this specification. Moreover, we find that birth control clinics led to a reduction of stillbirths by 1,000 population by 5.6 percent, of stillbirths by 1,000 births by 3.2 percent, and of infant mortality by 3.5 percent within ten years after establishment. These results suggest that birth control clinics particularly reduced births that posed significant health risks. Insignificant and small pre-treatment coefficients corroborate the validity of the common trends assumption. Using complementary city-level data, we do not find any significant effects of birth control clinics on puerperal deaths. The results are robust to extending the period of observation, using the last treated counties instead of never-treated counties as a control group, and employing the estimator proposed by Callaway and Sant'Anna (2021) instead of Sun and Abraham's (2021) interaction weighted estimator. Empirical evidence using data on all-age mortality and waterborne typhoid deaths suggests that the roll-out of birth control clinics is not confounded by more general improvements in counties' health infrastructure such as sewerage systems and waterworks. Moreover, we provide several arguments why the Sheppard-Towner Act and early welfare programs such as Mothers' Pensions are also unlikely to confound the estimates.

Our contribution to the literature is threefold. First, we add to the general debate about the role of family planning services (see, e.g. Udry *et al.*, 1976; Cutright and Jaffe, 1977; Molyneaux and Gertler, 2000; Mellor, 1998; Bailey, 2012, 2013; Canning and Schultz, 2012). In particular, we look at the first family planning initiatives in the history of the U.S. that have so far only been discussed qualitatively (see, e.g., Reed, 1978; Chesler, 1992; McCann, 1994; Gordon, 2002; Hajo, 2010; Engelman, 2011). Second, we add to the general understanding of the U.S. demographic transition by focusing on its final phase before birth rates started to increase again for roughly twenty years (see, e.g. Bourne Wahl, 1992; Steckel, 1992; Haines, 2000; Greenwood and Seshadri, 2002; Hacker, 2003; Haines and Hacker, 2006; Jones and Tertilt, 2006; Kitchens and Rodgers, 2023; Curtis White, 2008; Bleakley and Lange, 2009; Bailey and Collins, 2011; Guinnane, 2011; Wanamaker, 2012; Aaronson *et al.*, 2014; Hansen *et al.*, 2018; Bailey and Hershbein, 2018; Beach and Hanlon, 2023; Ager *et al.*, 2020; Grimm, 2021; Alsan and Goldin, 2019; Costa, 2015). At the same time, in our period of observation, the U.S. experienced a strong decline in infant and child mortality (see, e.g. Cutler *et al.*, 2006; Haines, 2006). Third, we complement the literature that focuses specifically on the role of anti-abortion legislation (Lahey, 2014a,b, 2022) and modern contraception (Goldin and Katz, 2002; Bailey, 2006, 2010; Bailey *et al.*, 2023) in the U.S. The situation in the early 20th century differs from the situation in the U.S. in the second half of the 20th century when the pill was introduced. People were poorly informed about contraceptives; even the medical profession was not trained in these issues. Moreover, contraceptives were difficult to find in pharmacies and other stores as supplies were unpredictable and modern highly effective contraceptives did not yet exist (Engelman, 2011). Still, studying early birth control clinics provides valuable insights into the fundamental role of contraceptives for socio-demographic change. Our results also have implications for low income countries today, as they demonstrate the power of family planning in a context where parents, and especially mothers, desire smaller family size and longer birth intervals but are constrained to implement these preferences. The study also indicates that family planning can have substantial indirect positive effects on health.

The remainder of the paper is structured as follows. In Section 2, we provide information on the historical context, the birth control movement, the roll-out of Sanger's birth control clinics and the services these clinics delivered. In Section 3, we introduce the data sets we use. Section 4 explains the empirical strategy and presents the estimated effects of birth control clinics on fertility using the Census data. Section 5 explains the empirical strategy we employ using the county level vital statistics, cross-validates the estimated effects on fertility and complements these results with an analysis of the effects on stillbirths, infant mortality and puerperal deaths. Section 6 concludes.

2. Historical Background

2.1. The U.S. fertility transition, the Comstock Act, and anti-abortion laws

The fertility transition in the U.S. started to take speed in the mid 19th century (Jones and Tertilt, 2006; Bailey and Hershbein, 2018). At the time, most couples used withdrawal and abstinence as well as extended breastfeeding to space births (Engelman, 2011; Reed, 1978; David *et al.*, 1986), but also abortion and—possibly, as a method of last resort—infanticide, yet data on infanticide is scarce (see, e.g., Wheeler, 2012). The second half of the 19th century also witnessed the emergence of the first rudimentary contraceptive devices such as condoms, vaginal sponges, douches, diaphragms and rubber syringes. Initially, these devices were widely advertised in newspapers and flyers and marketed by pharmacies, doctors, midwifes, druggists, and other entrepreneurs (Engelman, 2011). At the same time, marriage manuals diffused that explicitly considered that couples could and should enjoy sexual pleasure without procreation (see, e.g., Knowlton, 1832; Owen, 1876).

A Christian morality movement, led by Anthony Comstock, an influential political figure, disapproved this development and associated contraception with illicit sex and pornography. In 1873, Comstock succeeded in making the U.S. Congress codify the Comstock Act, which made it illegal to mail, ship, or import articles, drugs, medicines, and printed materials deemed "obscene", including any article or information related to the prevention of conception. As a consequence, 45 U.S. states passed or amended anti-obscenity statutes explicitly mentioning contraception with differences across states in the level of enforcement (Bailey, 2010). People not complying with the law could be subject to fines or even arrest, with the respective goods confiscated. Additionally, between 1860 and 1890, many states passed strict anti-abortion laws (Engelman, 2011; Lahey, 2014a; Lahey, 2014b). As one can imagine, it was difficult to enforce the Comstocks laws in any possible way. However, in particular for medical professionals strict Comstock laws meant that they could no longer sell or offer advice on contraceptives without risking their licenses (Bailey and Hershbein, 2018). Guttmacher (1947) shows that physicians with a medical degree earned before 1920 had typically not received any training on contraception at medical school. Hence, women had to rely on ill-informed sources and mostly ineffective contraceptives.

Despite these restrictions, the birth rate, i.e., births per 1,000 population, fell by more than 30 percent for both Whites and Blacks in our period of observation, i.e., from 29.2 in 1910 to 18.6 in 1940 for Whites, and from 38.5 in 1910 to 26.7 in 1940 for Blacks (Haines, 2006). The literature on the determinants of the fertility decline during the first half of the 20th century is still scare. Kitchens and Rodgers (2023) show that agricultural price increases explain about 9 percent of the overall decline in fertility from 1910 to 1930. Their estimates are consistent with agricultural booms increasing the opportunity cost of children for agricultural women with negative effects on fertility. For women who entered adulthood during the Depression in the 1930s, Bellou and Cardia (2021) and Bellou *et al.* (2024) show that shocks to household income resulted in an employment increase via an added-worker effect and a fertility decrease; their estimates imply that from 1930 to 1936, the Great Depression accounts for almost 40 percent of the fertility decline, which is in line with previous work from Fishback *et al.* (2007). In the late 1930s, the fertility decline slowed down or came to a halt, which can partly be explained by the positive fertility effects of the New Deal relief programs (Fishback *et al.*, 2007).

2.2. The birth control movement

In opposition to the restrictive interventions of the Comstock laws, a birth control movement and activism for sexual freedom emerged at the beginning of the 20th century.[1] Margaret Sanger, a nurse, whose mother had been through 18 pregnancies in 22 years and died at age 50 of tuberculosis and cervical cancer, became a leading figure of this movement. Sanger was concerned about the hardship brought about by repeated pregnancies, childbirth and self-induced abortions among poor women. Therefore, she devoted her life to make "birth control", a term she coined to refer to contraception, legal and widely accessible (McCann,

[1] McCann (1994) provides a detailed chronology of events in the U.S. birth control movement.

1994). During a visit to the Netherlands, Sanger was struck by the low rates of infant and maternal mortality, which were not least attributed to the availability of birth control clinics providing professional contraceptive counseling to women (McCann, 1994; Hajo, 2010).

In 1916, Margaret Sanger established the United States' first birth control clinic (or center for contraceptive instruction) in New York, Brownsville, Brooklyn. The clinic provided women with advice on effective contraception, sexual health and hygiene, and was sponsored by wealthy supporters. It was promoted through flyers in Yiddish, Italian, and English (see Appendix Figure A.1) as well as press releases (Reed, 1978). Sanger was convinced that "occlusive diaphragms", at a cost of one to two dollars, in conjunction with a spermicidal jelly were the most effective contraceptive method available to women at the time and that proper instruction and follow-up visits were crucial to ensure their efficacy. Indeed, the effectiveness of the diaphragm, if correctly applied, reached 90-95 percent and it had the advantage over withdrawal, condoms or periodic abstinence that it put women in control of contraceptive decisions (Hajo, 2010; Lane *et al.*, 1976).[2] McCann (1994) explains that a "setting was needed where trained practitioners could screen women for existing health problems, fit devices, and teach women to insert the device themselves." In the presence of strict Comstock laws, Sanger concluded that only (illegal) birth control clinics run by activists could provide such services (McCann, 1994; Hajo, 2010). Sanger was generally against abortions because of the associated health risks. Moreover, supporting abortions would have made the acceptance by the morality movement even more difficult (Engelman, 2011; Hajo, 2010). Therefore, birth control clinics should focus on contraceptives, in particular diaphragms, and refer women to doctors for therapeutic abortions only in exceptional cases. Despite the Brownsville clinic's success, with over 480 women attending in the first few days, it was shut down by the police ten days after it opened (Engelman, 2011; Hajo, 2010). Margaret Sanger was subsequently sentenced to 30 days in jail.

Sanger founded the Birth Control League of New York, which organized mailings, conferences, lectures as well as exhibits and advocated for legislative change. Similar leagues emerged in other cities, leading to the formation of over 30 birth control organizations across the country by 1917. In 1921, Sanger established the American Birth Control League; by 1924, it had more than 27,000 members and 10 branches. Sanger also founded and edited the Birth Control Review, a journal devoted to the birth control movement and mainly sold on the streets (Engelman, 2011). Sanger traveled extensively to give speeches, form alliances and raise funds for her cause. However, her speeches were often interrupted or halted by the police, and she faced multiple criminal accusations, resulting in short stints in jail. She also visited European countries to learn from their

[2] Douches were another available contraceptive device women were in control of, but they were clearly less effective and potentially dangerous (David *et al.*, 1986).

Figure 1 — The Sanger Clinic, 46 Amboy Street, Brooklyn

Notes: Photo published by the Social Press Association, New York, on October 27, 1916, provided by the Library of Congress Prints and Photographs Division Washington, D.C. The photo shows mothers waiting in front of the United States' first birth control clinic in 46 Amboy Street, Brooklyn, New York.

more advanced family planning policies. Over time, Sanger slightly adapted her strategy in her fight for the legalization of contraception. Instead of emphasizing women's right to self-determination, she began to highlight the health benefits of child spacing for both mothers and children. With this strategy, she hoped to gain support from doctors and hospitals while avoiding legal conflicts (McCann, 1994; Engelman, 2011). Her ultimate goal was to establish birth control clinics across the nation (McCann, 1994).

The nationwide roll-out of Sanger's birth control clinics commenced in the mid-1920s. The widespread prevalence of sexually transmitted diseases among U.S. servicemen during World War I had turned sexuality and contraception into a public health issue and thus a legitimate scientific research topic (Engelman, 2011). In the course of the 1930s, legal victories further strengthened the birth control movement and weakened the anti-contraception laws. Simultaneously, birth control became more and more accepted by the public. Although contraception was finally approved by the American Medical Association in 1937 (McCann, 1994), it was only tentatively endorsed by the medical profession. Moreover, advertising and sales bans on contraceptives persisted in many U.S. states until the end of the 1960s (Bailey, 2010). Therefore, women still often relied on ill-informed sources providing ineffective contraceptives with many of them depending on mens' cooperation (Tone, 2002), not comparable with the diaphragms provided by birth control clinics. There were two types of birth control clinics: independent clinics, such as the Brownsville clinic, typically operated by female activists and a part-time male medical doctor, often an activist's spouse (Hajo, 2010), and clinics within hospitals. Over time, the clinics gradually professionalized and were increasingly staffed with skilled practitioners who were either trained in the clinics themselves or in Sanger's Clinical Research Bureau in New York.[3] Birth control clinics provided their services to women regardless of their ability to pay.[4] The clinics typically admitted only married women who already had children since they primarily aimed to enable women to space births, rather than to prevent births altogether (Hajo, 2010). Initially, the clinics did not target Black women due to strict segregation laws. However, the Black community began to establish their own Sanger clinics following the same model. As segregation laws started to change, clinics served both White and Black women (Hajo, 2010). Over the entire period, Black women made up about 11 percent of all patients, which roughly corresponded to their population share (Hajo, 2010).

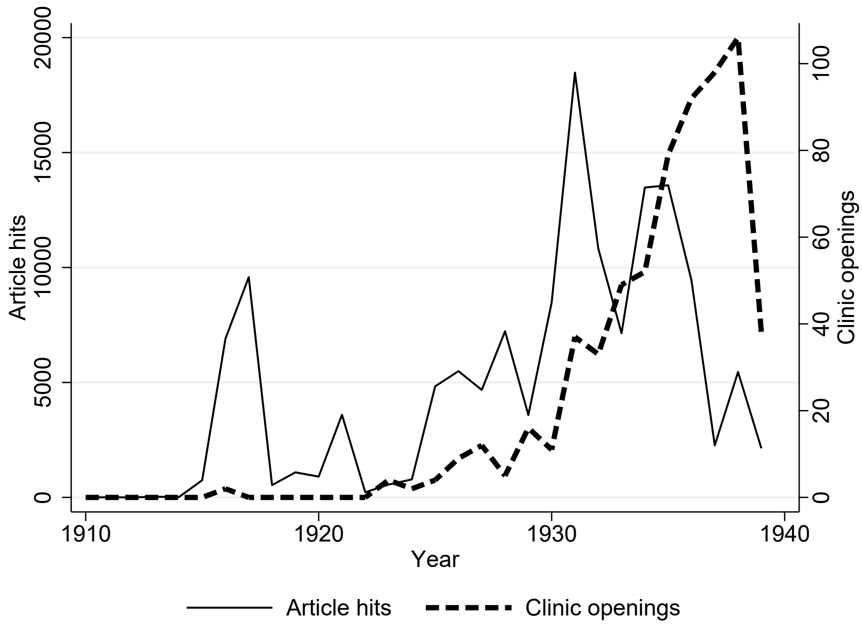
Between 1916 and 1940 more than 650 "Sanger clinics" were established in the U.S (Hajo, 2010). As shown in Figure 2, this process was prominently covered by the press. The figure plots the clinic openings (right scale) and newspaper articles mentioning the term "birth control" (left scale) over time. The first spike is clearly visible when the Brownsville clinic opened in 1916. In the following years, the spikes in the clinic expansion correlate with the spikes in press coverage. As opposition dwindled and birth control clinics became increasingly accepted in U.S. society in the 1930s, newspaper coverage declined and the correlation between newspaper coverage and the expansion of birth control clinics disappears in the data.

In 1939, the American Birth Control League merged with Sanger's Clinical Research Bureau to form the Birth Control Federation of America. At that point, Sanger became honorary president and relinquished active leadership. The organization changed its name to Planned Parenthood Federation of America in 1942. Margaret Sanger raised funds, among others from Katharine Dexter McCormick, to support the development of the pill. Her activism for women's rights, birth control and planned parenthood continues to receive praise today. However, she is also criticized for her associations with eugenics and support for forced sterilization (Weisbord, 1973; Engelman, 2011). Sanger died in Tucson, Arizona in 1966.

[3]The Clinical Research Bureau was the first legal birth control clinic in the United States, and soon grew into the world's leading contraceptive research center (Reed, 1978; Hajo, 2010).

[4]While the costs for an average patient was around USD 6.50, the highest fee charged was typically USD 5 (Reed, 1978).

Figure 2 — Birth control clinic openings and press coverage



Notes: Data source: Birth control clinics statistics by Hajo (2010) and Newspapers.com. The left y-axis shows the number of press articles using the term "birth control" by year. The right y-axis shows the number birth control clinic openings by year.

3. Data

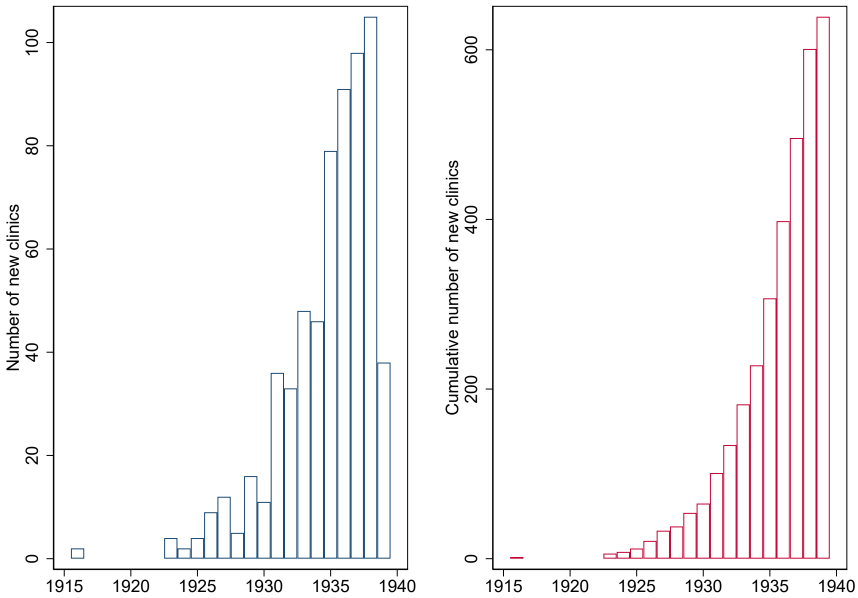
3.1. Birth control clinics

To measure exposure to a birth control clinic, we use a complete inventory of birth control clinics established in the U.S. before 1940. We compiled the data set by digitizing information gathered by Hajo (2010) from various sources, including historical issues of the Birth Control Review and press archives. For each clinic, we obtained data on the county in which it was located, the year of its establishment as well as, if applicable, the year of its closure. The data set encompasses a total of 639 birth control clinics, which are geographically dispersed across 44 states.[5]

Figure 3 gives a first impression of the roll-out of birth control clinics over time. The left panel shows the number of newly established birth control clinics by year, while the right panel shows the cumulative number of clinics. The first birth control clinic opened in Brooklyn in 1916; a second one in the same year in St. Paul, Minnesota. Further birth control clinics did not open until the year 1923. In the early to mid-1930s, we observe increased dynamics in the roll-out. By the end of 1939, 639 birth control clinics had been established in the U.S. Note that some clinics closed soon after opening, primarily due to financial instability or a lack of staff (McCann, 1994). In most of these cases, the closed

[5]We exclude the states of Alaska and Hawaii from the analysis.

Figure 3 — Birth control clinics over time

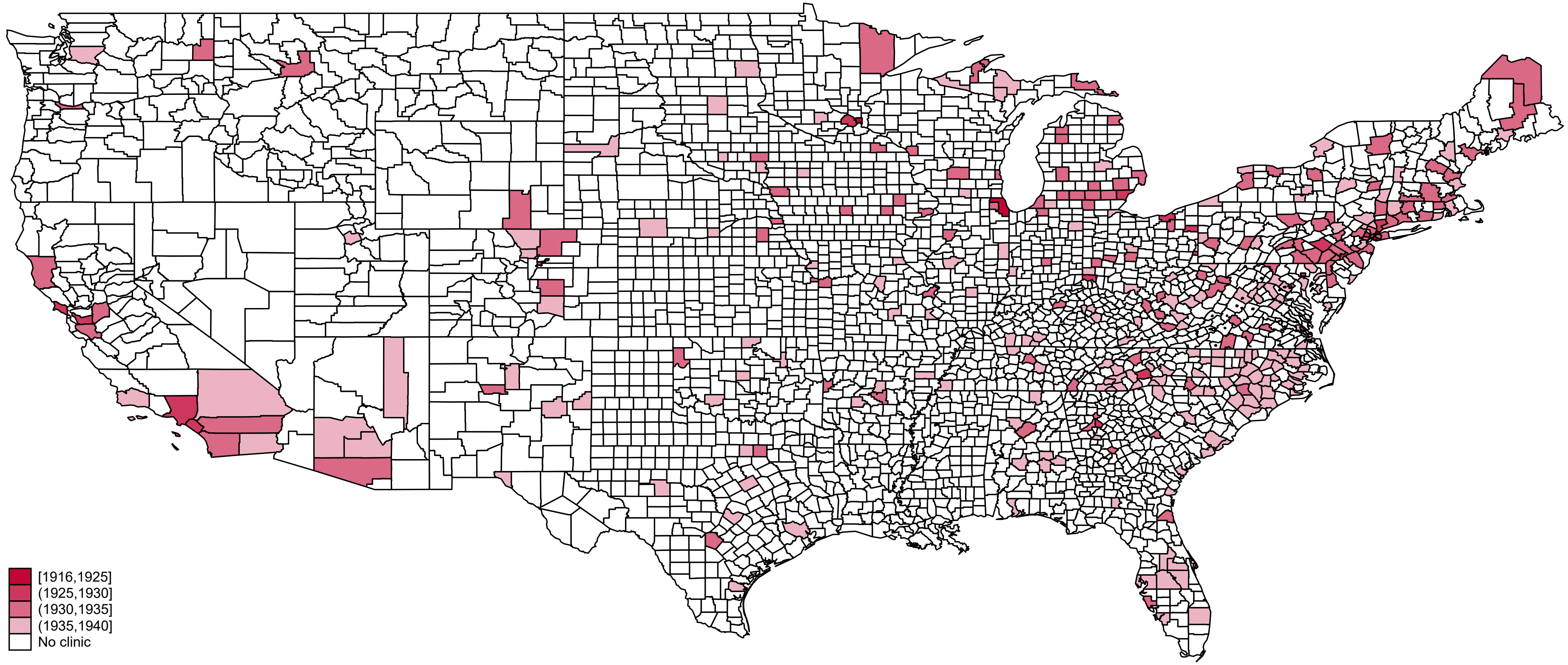


Notes: Data source: Birth control clinics statistics by Hajo (2010). The left panel shows the number of new birth control clinics by year. The right panel shows the cumulative number of new birth control clinics by year.

clinic was replaced by another clinic in the region shortly afterward, or other clinics in the region were still available. Therefore, we use the year of the opening of a county's first clinic as the treatment variable in the empirical analysis. This also respects the fact that once a clinic had been operating it could have had an enduring effect as diaphragms had been distributed, information had spread and norms might have changed.

It was not uncommon for women to travel to a clinic in a neighboring county if there was no clinic in their county of residence. For example, Engelman (2011) reports that women visiting the Brownsville clinic in Brooklyn came from all five boroughs of New York, Long Island, New Jersey and even as far as Philadelphia or New England. Therefore, in the empirical analysis, the variable measuring the presence of or the exposure to a birth control clinic takes the value one if a birth control clinic had opened in the county or any adjacent county before or in the respective year. While married women might have traveled across county borders to access a birth control clinic, the historical literature does not provide any evidence of families migrating because of birth control clinics. Once women had learned how to correctly insert a fitted diaphragm, they could have used it for years, making selective migration highly unlikely.

FIGURE 4 — Birth control clinics in U.S. counties 1916–1940



Notes: Data source: Birth control clinics statistics by Hajo (2010). The map shows U.S. counties. The earlier a county was exposed to a birth control clinic, the darker it is colored.

We illustrate the geographic dimension of the roll-out of birth control clinics in Figure 4. Counties that were exposed to a birth control clinic earlier are represented by darker shading. Initially, birth control clinics were established in larger cities such as New York, Chicago, and Los Angeles, before extending into smaller cities and rural areas. There is considerable heterogeneity even across counties within states. The establishment of clinics was not based on a specific roll-out plan; they were founded where Sanger could find volunteers to serve in these clinics and wealthy supporters to fund them. Especially in bigger cities, additional clinics often opened soon after the establishment of the first clinic (see Figure A.2 in the Appendix).

To further explore the roll-out, we run regressions that analyze the county level correlations between the establishment of a clinic and socio-economic county characteristics conditional on state fixed effects (see Table A.1 in the Appendix). We use both an indicator variable depicting whether a birth control clinic had been established in the county or an adjacent county by 1940, and a continuous variable measuring the number of years a clinic had existed in the county or an adjacent county by 1940 as outcome variables of these regressions. We start with a specification in which we use baseline fertility in 1920, urbanity indicators in 1920 (urban population share, population share living in farm households, population share living in big cities with >100.000 inhabitants) as well as the population age structure in 1920 (share of women aged 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, and 45-49) as predictor variables. Then, we add a set of covariates describing the socio-economic background of women of fertile age in 1920 such as the sex ratio in the population aged 15-49, the share of foreign born women, the share of literate women, the share of Black women, the share of non-White and non-Black women as well as the share of women in the labor force. Finally, we add a set of covariates further describing the socio-economic situation of the population such as the share of adults ever married, the population size (in 1,000), the Protestant share (reference category), the Catholic share and the other religion share, an occupational income score as well as the manufacturing employment share. We also account for the share of World War I veterans among male adults, as veterans often returned with sexually transmitted diseases which turned sexuality and contraception into a public health issue and might have supported the establishment of a clinic. In line with the historical literature, we find that birth control clinics were first established in big cities and only later spread into more rural counties. Therefore, we control not just for county fixed effects but also for an urban location of the household in all Census regressions. Most other socio-demographic variables fail to predict if or when a birth control clinic was established. The only other variable that is consistently correlated with the clinic roll-out across all specifications is the share of women in the labor force in 1920. To investigate whether this poses a thread to the validity of our empirical approach, we use female labor force participation as the outcome variable of an event-study along the lines of Equation 2. The pre-treatment coefficients are all close to zero and insignificant, which suggests that birth control clincis did not open in areas where female labor force participation was on the rise (and fertility potentially on the decline) (see Appendix Figure A.3). This finding is reassuring since in our empirical analysis, we assume that when exactly a birth control clinic was established in a county is as good as random. We present a battery of additional checks to validate this assumption. For example, we thoroughly investigate pre-trends in our event study design and find that there is no conspicuous change in a county's birth rate in the years before the birth control clinic was established. In another check, we include state by cohort fixed effects to control for potentially differential fertility trends across states. Moreover, we include baseline fertility and baseline female labor force participation interacted with year fixed effects to account for differential fertility trends across counties with different baseline fertility or baseline female labor force participation levels. In extended analysis, we restrict the sample to big city counties to further enhance comparability. These exercises univocally support the validity of our approach.

3.2. Census data

To assess the impact of the exposure to a birth control clinic at the individual level, we use information on women from the Population Census of 1920, 1930 and 1940. Harmonized full count data is available from the Integrated Public Use Microdata Series (IPUMS USA) (Ruggles *et al.*, 2022). We exclude the states of Alaska and Hawaii as well as members of the armed forces residing in training camps or in other military institutions. Since the data provide information on the respondents' county of residence, we can merge the birth control clinic data to the Census data at the county level. We use a time-constant set of counties throughout our period of observation employing the county longitudinal template by Horan and Hargis (1995) with 1920 as the base year. Thereby, we avoid that changing county borders over time blur our analyses.[6] We limit the Census sample to White and Black women as the remaining share of women is small (0.35 percent) and heterogeneous. Moreover, we focus on marital fertility since birth control clinics did not target unmarried, childless women.[7]

To measure fertility, we use the number of a woman's own children below the age of five living in the same household.[8] In robustness checks, we alternatively use the number of all children of a woman living in the same household without any age restriction for

[6] In our period of observation, there were only few changes to the borders of counties in the United States; most changes happened earlier in American history. Reasons for changes in counties were mostly the consolidation of counties and the reorganization and creation of new counties. Applying the Horan-Hargis longitudinal template to our data leads to a reduction in the number of "counties" by 67. This mainly results from grouping counties to have consistent borders over time.

[7] Giving birth outside marriage was rare at that time; moreover, out-of-wedlock births were not well-accepted by society, which may result in severe reporting error.

[8] Note that the Census variable explicitly refers to a mother's own children; thus, multi-family households or siblings do not pose a problem for this fertility measure.

the children. We limit the sample to women aged 15 to 39 and exclude older women as the likelihood that women above the age of 40 have children below the age of five at home is low in our period of observation (Grimm, 2021). The number of children ever born is not available in the full-count Census of the years 1920 and 1930; it is available but vastly incomplete in the 1940 Census as only five percent of all women were asked about their children ever born.[9] Relying on retrospective fertility in 1940 would come with two additional problems we would like to minimize, namely selective mortality as well as migration across counties. The latter would result in assigning women to counties they did not live in during their fertile period; as a result, assigned exposure to birth control clinics might suffer from measurement error. All analyses using the Census data are executed at the individual, i.e., the woman, level.

The Census data further provide us with a set of observable individual characteristics that we use as control variables. We use information on a woman's age and race, her literacy status[10], an indicator for whether she was born in the U.S., whether she lives in a farm household, and whether the household is located in a rural or urban area, i.e., typically in a town of more than 2,500 inhabitants. Moreover, we use an indicator for whether the household is located in a big city of more than 100,000 inhabitants.[11]

Table 1 shows descriptive statistics for the total sample, and separately for the 1920, 1930, and 1940 Census samples. In total, we observe more than 45 million women. The average woman in our sample is exposed to a birth control clinic for 2.745 years (including zeros); this number increases from 0.246 in 1920 to 5.773 in 1940. The average number of a woman's children below the age of five is 0.629; it is highest in 1920 (0.747) and decreases to 0.541 in 1940. Women are on average 29 years old, 11 percent are Black, and also 11 percent are foreign-born. The literacy rate is as high as 96 percent. 22 percent of women live in farm households, 57 percent reside in urban areas, and 30 percent in big cities.

3.3. County level vital statistics

As a complementary data set, we use county-level natality and mortality data compiled by Bailey *et al.* (2016). The team has digitized historical print sources from 1915 onward, double-checked the entries for accuracy, and processed the data to get consistent information at the county level. While the share of reporting counties was still low in 1915, it increased over the years; from 1933 on, the data set includes vital statistics from all counties. We

[9] https://usa.ipums.org/usa/chapter2/chapter2.shtml.

[10] In the 1940 Census, literacy status is missing. Therefore, we use information on education to proxy literacy status. Any woman who reports to have achieved at least grade 4 is considered to be literate.

[11] The Census data reveal there were 286 cities with a population of more than 25,000 inhabitants in 1920. 68 or 23.78 percent of them had a population of more than 100,000. In 1930, there were 365 cities of which 93 or 25.48 percent had a population of more than 100,000. In 1940, there were 398 cities of which 92 or 23.12 percent had a population of more than 100,000.

TABLE 1 — Descriptive statistics: Census data

	1920 (I)	1930 (II)	1940 (III)	All years (IV)
Years of exposure to a BCC	0.246	1.550	5.773	2.745
(St. Dev.)	*(0.959)*	*(3.369)*	*(5.437)*	*(4.562)*
# childr. <5 in HH	0.747	0.626	0.541	0.629
(St. Dev.)	*(0.876)*	*(0.822)*	*(0.767)*	*(0.823)*
Age	29.333	29.518	29.544	29.474
(St. Dev.)	*(5.910)*	*(5.984)*	*(5.890)*	*(5.928)*
Black (=1)	0.101	0.122	0.113	0.113
Foreign born (=1)	0.172	0.121	0.064	0.114
Literate (=1)	0.935	0.966	0.969	0.958
Protestant†	0.522	0.517	0.524	0.521
Catholic†	0.307	0.311	0.306	0.308
Other religion†	0.172	0.173	0.171	0.172
Farm hh (=1)	0.256	0.210	0.199	0.219
Urban residence (=1)	0.543	0.598	0.577	0.574
Big city residence (=1)	0.278	0.320	0.295	0.299
Obs. (women)	13,001,035	15,336,423	16,783,281	45,120,739

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940 and NHGIS. The table shows the means of the variables in the total sample, and separately for the Census years 1920, 1930, and 1940. Standard deviations of non-binary individual-level variables are provided in parentheses. †Religious composition is based on county level information, numbers show the average of population weighted county shares.

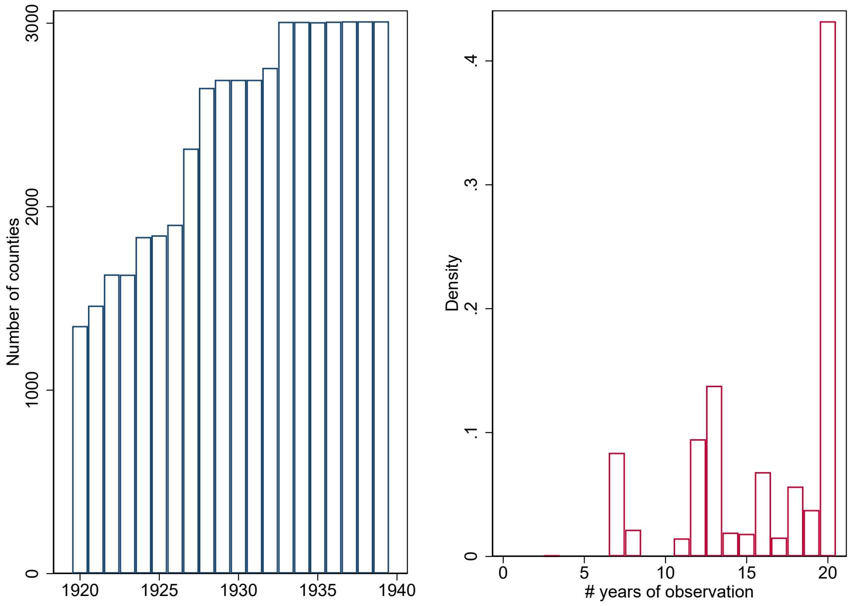
merge these data to the birth control clinic data at the county-year level; again, we employ the county longitudinal template by Horan and Hargis (1995) with 1920 as the base year to create a panel of consistently observed counties or county groups over time.

In our main analysis, we use annual information on the number of live births (exclusive of stillbirths) and infant deaths (i.e., deaths of children below the age of one exclusive of stillbirths). In a validity check, we use the total number of deaths by all ages and subtract infant deaths to obtain a mortality measure unaffected by infant deaths. Births and deaths are assigned to the county of occurrence in this period of observation. To account for changes in counties' population, we compute yearly crude birth and mortality rates by dividing the number of live births and the number of deaths in a year by the county's total population in the respective year.[12] For infant deaths, we use the number of births instead of the total population as the denominator.

Additionally, we merge county-level data from IPUMS NHGIS (Manson *et al.*, 2022) on the yearly number of stillbirths by place of occurrence in the period from 1922 to 1939 to the vital statistics collected by Bailey *et al.* (2016). To obtain stillbirth rates, we divide the number of stillbirths by total population, and, alternatively, by the sum of stillbirths

[12]Since female population between ages 15 and 44 is not available in the data before 1930, we cannot compute birth rates as the number of births divided by the number of women of child-bearing age.

FIGURE 5 — Overview of vital statistics



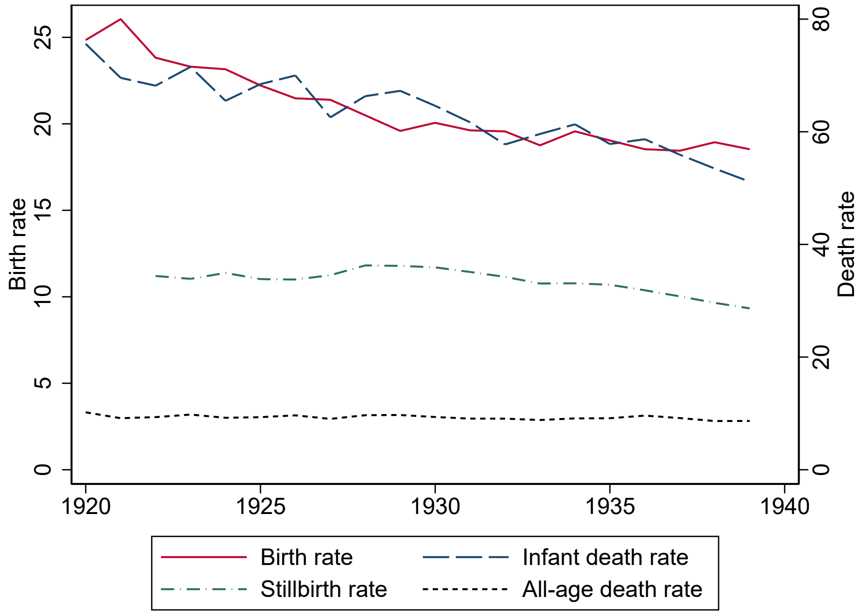
Notes: Data source: U.S. Vital Statistics. The left panel of the figure shows the number of counties available in the data set by year. The right panel shows the density of the number of observations by county in the period from 1920 to 1939.

and live births. Thus, we account for changes in fertility and ensure that the stillbirth rate does not just decrease because the number of pregnancies decreases.

Figure 5 provides an overview of the number of observations available in the vital statistics data set. The left panel shows the number of counties in the data set by year. In 1920, we have information from 1,350 counties. This number steadily increases over the following years. From 1933 on, the data set includes observations from more than 3,000 counties. The right panel shows the distribution over the number of observations per county. We observe 43 percent of all counties every single year in the period from 1920 to 1939. For 63 percent of all counties, we have observations from 15 years at least.

The development of crude birth and mortality rates in our period of observation is depicted in Figure 6. The number of births decreases from 25 per 1,000 population in 1920 to 19 per 1,000 population in 1939. In the same period, the number of infant deaths decreases from 76 per 1,000 births to 51 per 1,000 births. Thus, we observe a 24 percent decrease in the crude birth rate and a 33 percent decrease in the infant death rate within twenty years. The number of stillbirths per 1,000 births decreases from 34 in 1922 to 29 in 1939, which is a decline of 15 percent. The all-age mortality rate (without infant deaths) only slightly decreases from around 10 to 9 per 1,000 population within the period of observation. Further descriptives on these vital statistics can be found in Appendix Table A.8.

FIGURE 6 — Birth and death rates in the vital statistics



Notes: Data source: U.S. Vital Statistics. The figure shows birth rates (births per 1,000 population), stillbirth rates (stillbirths per 1,000 births incl. stillbirths), infant death rates (infant deaths per 1,000 births), and all-age death rates (deaths without infant deaths per 1,000 population) over the period from 1920 to 1939.

3.4. City-level causes of deaths data

To investigate the impact of birth control clinics on maternal health, we additionally use city-level data on causes of deaths. The U.S. Census Bureau has collected annual city-level mortality data and published them in the Vital Statistics of the United States annual volumes since 1900. Ager *et al.* (2024) have digitized the causes of death data collected by the U.S. Census Bureau, merged this data with the full-count Population Censuses collapsed at the city level, and provided us with access to the data set. Similar to the county-level data, the city-level data do not report deaths by gender in our period of observation; however, they do report deaths by specific causes. This allows us to focus on puerperal deaths, which are most closely associated with potential maternal health effects from birth control clinics. Puerperal deaths refer to death events due to postpartum bacterial infections of the female reproductive tract following childbirth or miscarriage; they are also known as childbed fever and puerperal fever deaths. To account for the size of the underlying population, we divide the number of puerperal deaths by the female population aged 15 to 49. We use a linear interpolation to estimate the female population in each city in intercensal years. In a validity check, we use the number of deaths due to typhoid fever and divide them by the city's total population to obtain death rates in an analogous way.

We restrict the sample to the period of 1920 to 1937. Following Ager *et al.* (2024), we stop in 1937 because the coding of causes of death significantly changes after 1937 (Feigenbaum *et al.*, 2019). We keep cities in our sample if they consistently report causes of deaths from 1920 to 1937. Following this procedure, we can use data from 315 cities. We merge the birth control clinic data to the causes of death data at the county level. Descriptives on the city-level causes of death data can be found in Appendix Table A.8.

3.5. Further data

We use Census data to construct a set of additional county level variables and use them as predictor variables in roll-out regressions, in robustness tests, heterogeneity analyses, or in a matching regression. Based on 1920 Census data, we construct urbanity indicators (urban population share, population share living in farm households, population share living in big cities with >100.000 inhabitants), age structure variables (the share of women aged 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, and 45-49 in the population), variables describing the socio-economic background of women of fertile aged 15-49 (the sex ratio, the share of foreign born women, the share of literate women, the share of black women, the share of non-white and non-black women as well as the share of women in the labor force), and further variables describing the socio-economic situation of a county's overall population (share of adults ever married and population size (in 1,000)). Additionally, we compute county-level occupational income scores; the occupational income score in the Census assigns each occupation in all years a value representing the median total income (in hundreds of 1950 dollars) of all persons with that particular occupation in 1950. Based on 1930 Census data, we construct a county level variable measuring the share of World War I veterans in the male population aged 18 and older.

Data from IPUMS NHGIS (Manson *et al.*, 2022) allow us to measure a county's religious composition in 1926, i.e., the share of Catholics, Protestants, and followers of another religion, the share of manufacturing employment in 1930, and the share of unemployed workers in 1930. Moreover, we use data on voting outcomes of presidential elections of the years 1920 and 1932 from Robinson (1934). Finally, we use county level retail sales data from the years 1929 and 1939 provided by Fishback *et al.* (2005) as a proxy for the local intensity of the Great Depression as well as average per capita public works and relief spending in the periods 1933-1935 and 1933-1939 also provided by Fishback *et al.* (2005) to measure the local prevalence of social welfare programs. Appendix section A.2 gives an overview of all data sources and variables we use.

4. Evidence from Census data

4.1. Identification strategy

To identify the impact of a woman's exposure to a birth control clinic on fertility in the Census data, we estimate the following regression equation:

$$y_{ict} = \beta yexp_{ict}^{BCC} + x'_{ict}\gamma + \lambda_c + \delta_t + u_{ict}, \quad (1)$$

where y_{ict} is the number of a woman i's own children below the age of five living in the household in county c observed in Census year t. $yexp_{ict}^{BCC}$ is the number of years a woman has been exposed to a birth control clinic in her own or an adjacent county since her 15th birthday. Note that this measure varies by county and year but also within a county-year cell across different ages of women. The matrix of control variables x'_{ict} includes age fixed effects to flexibly control for fertility differences within the fertile period. It also includes an indicator for living in an urban area to account for the fact that fertility might differ between urban and rural areas while the roll-out of birth control clinics started in urban areas. Moreover, it includes a woman's literacy status, race, an indicator for being foreign born, an indicator for living in a big city, an indicator for living in a farm household, and the religious composition of the county where a woman resides. We include county fixed effects λ_c to capture unobserved time-invariant heterogeneity across counties. Survey year fixed effects δ_t account for general changes in fertility over time that are the same across counties. In our preferred specification, we interact the urbanity control with Census year indicators. Thereby, we allow fertility to have differential time trends in urban and rural areas. Standard errors u_{ict} are clustered at the county level.

The coefficient of interest β measures the impact of one year of exposure to a birth control clinic under the assumption that, conditional on county and Census year fixed effects and the set of controls, the years of exposure to a birth control clinic are orthogonal to unobserved determinants of fertility. We exploit variation from differences in the exposure to a birth control clinic across cohorts within the same county while flexibly controlling for general time effects that might differ between urban and rural areas, and general lifecycle effects, i.e., age fixed effects. Thus, the assumption can be reformulated in a way that resembles the standard parallel trends assumption for difference-in-differences designs: we assume that, in absence of exposure to a birth control clinic, women living in counties with and without a birth control clinic would have followed the same fertility trend conditional on our set of controls.[13]

To assess the validity of this key identifying assumption, we augment our analysis with

[13]To get a better impression of the changes in exposure over time, Figure A.4 in the Appendix shows for each Census year and for each cohort the mean exposure measured by the share of women exposed (upper figure), and by the mean years of exposure (lower figure). As can be seen, exposure increases over time and is higher for younger than for older cohorts.

an event study approach as laid out in the following equation:

$$y_{ict} = \sum_{\tau=-20}^{20} \beta_\tau exp_{ict,T+\tau}^{BCC} + x'_{ict}\gamma + \lambda_c + \delta_t + u_{ict}. \quad (2)$$

$exp_{ict,T+\tau}^{BCC}$ takes the value one if a woman i in county c and Census year t is observed τ years prior to or after the first-time exposure. All other variables are defined as in Equation 1 with the exception that we control for age and age squared instead of age fixed effects. Consequently, we compare women in the same county observed at different exposure times while controlling for general time as well lifecylce effects. In our preferred specification, we use intervals of three years instead of single years to increase statistical power. The first post-treatment interval runs from $\tau = 1$ to $\tau = 3$. Thereby, we take into account that a prevented pregnancy needs at least nine months to be reflected in the data. In the regressions, we omit the interval running from $\tau = -2$ to $\tau = 0$, i.e., the category just before treatment. Thus, the estimate for the lag category $\tau = 7$ to $\tau = 9$ compares women who have experienced an exposure time of seven to nine years at the time of the interview to women who are interviewed zero to two years before they get exposed for the first time (the omitted category). Again, standard errors u_{ict} are clustered at the county level.

If $\beta_{T+\tau}$ is zero for all $\tau < 0$, this supports the validity of a key assumption we make, namely that women living in treatment and control counties follow the same fertility trend in absence of the treatment, conditional on the set of controls. Apart from allowing us to investigate pre-treatment trends, the event-study approach also allows us to trace treatment dynamics over the years after the start of the exposure. For example, if the negative impact on fertility of a birth control clinic just slowly unfolds over time, we should observe that (negative) $\beta_{T+\tau}$ increases as $\tau \geq 0$ increases.

We check that our results are robust to changes in the empirical specification. In particular, we estimate a model with birth cohort by state fixed effects instead of county and Census year fixed effects. In additional robustness tests, we add interactions of Census year fixed effects and several other county variables such as the baseline fertility rate, the veteran population share, or proxies for the intensity of the Great Depression. Moreover, we restrict the sample to women living in big city counties to further enhance the comparability of treatment and control group observations.

4.2. Main results

We start our empirical analysis with a simple model, in which we regress the number of a woman's children below the age of five on state-specific cohort fixed effects while controlling for age fixed effects and for urban residence. Note that in this specification, we exploit variation across counties within the same state and cohort for identification. Controlling for a woman's age is indispensable despite the birth cohort fixed effects since the number of children naturally depends on a woman's age and, at the same time, older women are more likely to have lived more years exposed to a birth control clinic. Similarly, controlling for urbanity is essential since birth control clinics were initially established in urban areas before spreading to more rural areas, while living conditions in urban and rural areas do differ. Column 1 of Table 2 shows a highly significant negative effect of exposure to a birth control clinic on the number of woman's children below the age of five in the household.

In column 2 of Table 2, we move towards the county fixed effects specification of Equation 1, where the identifying variation comes from differences in the exposure to a birth control clinic across cohorts within the same county. We start with a stripped down version, in which we add only age fixed effects and the urban residence indicator to the county and Census year fixed effects. Again, we find a highly significant negative effect of years of exposure to a birth control clinic on the fertility measure. As compared to the state-specific cohort fixed effects model of column 1, the estimated negative effect has even slightly increased. In column 3, we add the full set of individual and county level socio-economic controls. While the point coefficient slightly decreases, we still find sizable and highly significant negative effects of exposure to a birth control clinic. In our preferred specification of column 4, we add an interaction of the urban residence indicator and Census year fixed effects to allow for differential fertility trends across urban and rural areas. The point estimate slightly increases. The estimated effect is not only statistically highly significant and but also economically meaningful: being exposed to a birth control clinic throughout the entire fertile period from age 15 to 39 reduces the number of a woman's children below the age of five in the household by about 16 percent ($\frac{25 \times 0.403/100}{0.628} \approx 16$). The average exposure time of women at the end of our observation period was 5.77 years, implying an average fertility reduction of 3.7 percent. Overall, exposure to birth control clinics can explain 5 percent of the total decline in the number of a woman's children below the age of five in the household in this period of observation.[14]

In a next step, we move to the event-study specification of Equation 2 to inspect pre-treatment trends and post-treatment dynamics. From Figure 7, we obtain three insights. First, the analysis provides evidence for the validity of the common trends assumption. The coefficients for the time periods prior to the exposure to a birth control clinic are insignificant and close to zero. Moreover, the overall pattern does not suggest any conspicuous systematic fertility trajectory. In particular, we do not find any evidence for fertility starting to decline already prior to the exposure to a birth control clinic. Secondly, we observe an immediate significant reduction in fertility once women are exposed to a

[14] Using the point estimate in column 4 of Table 2 (-0.403/100) multiplied by the average exposure time in our sample of 2.745 years yields a change in our fertility measure of -0.0111. The total change over the period 1920 to 1940 is -0.2063; hence, birth control clinics explain 5.36 percent of the total change. Birth rates in the US started to stagnate and then to increase in the mid-1930s (Bailey and Hershbein, 2018). Our estimates imply that exposure to birth control clinics slowed down this process.

Table 2 — The impact of exposure to a birth control clinic on fertility

	(I)	(II)	(III)	(IV)
Years of exposure to BCC (coef. × 100)	-0.347***	-0.389***	-0.276***	-0.403***
	(0.084)	(0.061)	(0.048)	(0.048)
Age FE	yes	yes	yes	yes
Urbanity control	yes	yes	yes	yes
Year FE		yes	yes	yes
State-specific cohort FE	yes			
County FE		yes	yes	yes
Socio-economic controls			yes	yes
Year FE × Urbanity				yes
Sample mean of dep. var.	0.628			
R^2	0.076	0.082	0.093	0.093
Observations	45,120,737	45,120,737	45,120,737	45,120,737

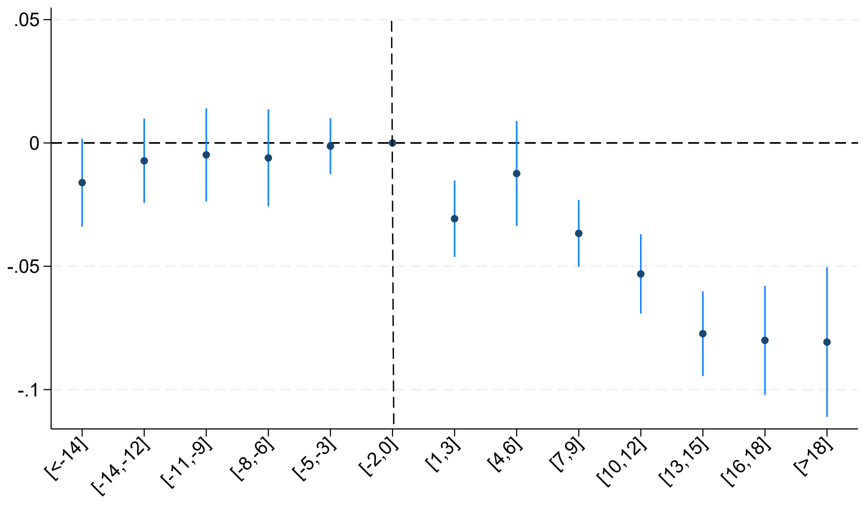
Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. The table shows OLS regressions. The dependent variable is the number of a woman's own children below the age of five living in the household. For readability, the coefficient associated with the years of exposure to a birth control clinic is muliplied by 100. Socio-economic controls are the literacy status, race, an indicator for being foreign born, an indicator for living in a big city, an indicator for living in a farm household, and the county's religuous composition. Standard errors are clustered at county level.

birth control clinic. Thirdly, this estimated negative effect is persistent and grows larger over the post-treatment years. Because only relatively few women are exposed to a birth control clinic for sixteen years or more, the confidence intervals for these long-term lag coefficients become larger. The increasing effect over time may partly be driven by further clinics emerging in the region after the establishment of the first clinic (see Figure A.2 in the Appendix). Note that the results from this event-study specification align well with the results from the simple county fixed effects specification above. For example, the event-study coefficient for the post-treatment interval $\tau = 7$ to $\tau = 9$ is -0.0367, which means that being exposed to a birth control clinic for 7 to 9 years (instead of 0) decreases the number of a woman's children below the age of five in the household by 0.0367. This is reasonably close to the result we obtain if we use the point estimate from our preferred specification of column 4 in Table 2, assume linearity and multiply it by 8 (mean of 7-9 years) and then divide it by 100 (since the coefficients in Table 2 were multiplied by 100): -0.403*8/100=-0.0322.

4.3. Robustness tests

To rule out the possibility that our findings reflect an increase in interpregnancy intervals without any effects on completed fertility, we perform a set of robustness checks. If birth control clinics just led to a postponement of births without any effects on completed fertility, we should expect negative fertility effects for younger women and positive fertility effects for older women. We estimate the model separately for different age groups of

Figure 7 — Event-study plots of the impact of exposure to a birth control clinic on fertility



Notes: Data source: IPUMS US Census, 1920, 1930 and 1940. The figure shows event study estimates of the exposure to a birth control clinic on the number of a woman's own children below the age of five in a household. Treatment effects are estimated along the lines of equation 2. Standard errors are clustered at the county level. The whiskers mark the 95 percent confidence band.

women (15-19, 20-24, 25-29, 30-34, 35-40). Our analyses reveal negative fertility effects throughout the entire age distribution (see Appendix Table A.2). The absolute effects are strongest for women aged 20-24 and then become weaker with age. Relative to the group specific means, the effects in general become weaker with age but slightly stronger again for the last age group. These results make it unlikely that birth control clinics just led to a postponement of births without any negative effects on completed fertility. To further alleviate this concern, we use Census information on the total number of a woman's own children living in the household (without any age restriction) as an alternative outcome variable. Estimating the model using this fertility outcome, we find similar results as in our main specification (see Appendix Table A.2).[15] Moreover, we check that the decreases in fertility are indeed driven by mothers averting higher-order births and not by women putting off motherhood completely. To this end, we create dichotomous outcomes variables that indicate whether a woman had "no child at age 30", or "no child at age 39". We do not find any significant effects of birth control clinics for these outcome variables (see Appendix Table A.2). Together, these results suggest that birth control clinics do not make women put off motherhood. Rather, women increase spacing between births and thereby reduce the number of births.

[15]Since this alternative outcome variable comes with the downside that it might be affected by selective moves out of the household due to fostering, schooling, vocational training or work, we decided to stick to the number of a woman's own children below the age of five in the household as our preferred outcome variable. Thereby, we focus on children at an age where they typically live with their mother and avoid any bias due to selective moves.

Table 3 — The impact of exposure to a birth control clinic on fertility - robustness checks

	(I)	(II)	(III)	(IV)	(V)	(VI)
Years of exposure to BCC (coef. x 100)	-0.529***	-0.465***	-0.453***	-0.454***	-0.354***	-0.448***
	(0.047)	(0.055)	(0.052)	(0.058)	(0.047)	(0.046)
Age FE	yes	yes	yes	yes	yes	yes
Urbanity control	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
County FE	yes	yes	yes	yes	yes	yes
Year FE × Urbanity	yes	yes	yes	yes	yes	yes
Socio-economic controls	yes	yes	yes	yes	yes	yes
County level fertility 1920 × year FE	yes					
Female labor force particip. in 1920 × year FE		yes				
Above median retail sales growth (1929-39) × year FE			yes			
Above median growth of social spending (1929-39) × year FE				yes		
Unemployment in 1930 × year FE					yes	
WWI veteran share in 1930 × year FE						yes
Sample mean of dep. var.			0.628			
R^2	0.094	0.093	0.092	0.092	0.093	0.093
Observations	45,120,737	44,917,303	41,658,206	41,658,206	45,120,737	45,062,891

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. The table shows OLS regressions. The dependent variable is the number of a woman's own children below the age of five living in the household. For readability, the coefficient associated with the years of exposure to a birth control clinic is muliplied by 100. Socio-economic controls are the literacy status, race, an indicator for being foreign born, an indicator for living in a big city, an indicator for living in a farm household, and the county's religuous composition. Standard errors are clustered at county level.

In another robustness test, we use our preferred specification, include further county specific covariates and interact them with Census year fixed effects. Thereby, we capture differential fertility trends depending on a county's baseline fertility level, baseline female labor force participation, and the local intensity of the Great Depression. To proxy for the local intensity of the Great Depression and its repercussions, we draw on Fishback *et al.* (2005) and use local retail sales growth in the 1930s, the local growth of social spending in the 1930s as well as the local unemployment rate in 1930. Moreover, we interact the population share of war veterans with Census year fixed effects, which allows us to flexibly control for war induced changes in the marriage market. Table 3 shows that these additional controls hardly affect the estimated impact of birth control clinics. Thus, we conclude that we do not find any evidence that differences in fertility or female employment prior to the roll-out of birth control clinics, the prevalence of war veterans, or the local intensity of the Great Depression confound our estimates.

Further subsample estimates provide evidence that the estimated effects are not confounded by the expansion of for-profit providers of birth control devices and irregular birth control clinics run by nurses, chiropractors and entrepreneurs. Despite the Comstock laws that continued to ban birth control, the irregular market for contraceptives expanded quickly in the 1930s, especially in some of the bigger cities such as Chicago (Holz, 2012). This market was fueled to a large extent by "charlatans, quacks" (p.46) and entrepreneurs without any medical training. They were primarily motivated by profit rather than a desire to promote charitable causes. The entrepreneurs often closely collaborated with the illegal manufacturers of contraceptives. Their product offerings did typically not include effective diaphragms (and certainly no professional advice how to fit them) that women could control but rather condoms, spermicides and contraceptive gels, most of it not very effective or dependent on men's cooperation. Still, to rule out that this rising commercialization of birth control confounds the estimates, we drop the 1940 Census and re-estimate the model only using the 1920 and 1930 Censuses, i.e., data from a period that preceded the rapid commercialization of birth control. The effect turns out to be even larger in this subsample (see Appendix Table A.3). Hence, we conclude that the estimated birth control clinic effects are not confounded by the development of the commercial, and largely illegal, market for contraceptives.

To address the concern that the estimated effects depend on the particular definition of the treatment, we re-define the treatment variable in additional robustness tests. In particular, we use log transformed exposure time since age 15 and sine transformed exposure time since age 15 as alternative treatment variables. One might criticize that measuring exposure since age 15 does not capture the "right" period in life where birth control clinics become relevant. To align the exposure measure more closely to the admittance criteria of birth control clinics ("married women who already had children"), we use Census data to compute the county-specific average age at first marriage for married women in the age group 25-34 in 1930. Then, we derive exposure time for each married woman in our data set by assessing how many years a woman had access to a birth control clinic from the average age at marriage in her county until her age at interview.[16]. Our findings are robust to all these different parameterizations of the treatment variable (see Appendix Table A.4).

Instead of relying on the simple rule of thumb that a clinic is relevant if it is in the county of residence or in any adjacent county, we model exposure to a clinic according to the availability of clinics in buffer zones around the centroid of a woman's county of residence. We consider a buffer zone of 20km, a buffer zone of 50km and a ring of 50 to 100km. The estimated effects within the buffer zones of 20km and 50km are very similar to the previous estimates (see Appendix Table A.5). Remember that women often also came to the clinics from adjacent counties to get birth control services. Combined with the fact that a 50km buffer zone largely corresponds to a county and its adjacent counties as the treatment unit, it is not too surprising to find similar effects for the 20km and the 50km buffer zone.[17] Moreover, the estimated effects are clearly larger within the buffer zones than in the ring around them. Note that counties are too small to allow for larger buffer zones and rings since the ring of a given county quickly intersects with buffer zones of many other neighboring counties, which blurs the analysis. Taken together, this exercise justifies the choice of restricting exposure to birth control clinics in the county of residence or adjacent counties. This choice is also preferred because the Census data do not provide exact geocodes of households, nor can we identify exact geocodes of the clinics in our inventory.

4.4. Heterogeneity

In additional analyses, we investigate whether the estimated average effect on fertility varies across socio-demographic groups, voting patterns, state groups, with the local intensity of the Great Depression or with the share of veterans in a county. We find that the impact for Black women is significantly larger than for White women, which might be due to a higher demand for birth control services. Note that these differences in the estimated effects for Black and White women are likely even understated since, in particular in earlier years, Black women might not have had actual access to birth control clinics due to segregation laws although they are coded as being exposed if they lived in a region with a birth control clinic. Furthermore, we observe highly significant negative effects both within big city counties and outside of big city counties, with the effects being larger outside (also relative to the respective group means). Since the roll-out of

[16] The full count Census provides information on the age at marriage for 90 percent of all married women in 1930 and 5 percent of all married women in 1940

[17] The average radius of counties is between 21 and 24km in the Northeastern, Midwestern and Southern states, while it is 52km in the Western states. Hence, a county plus the adjacent counties fit well within a radius of 50km.

birth control clinics started in big cities and only later spread to smaller cities and rural areas, it is reassuring to find sizeable and significant negative effects in both groups. The estimated effects do not differ significantly between women born in the U.S. and first-generation migrants or between counties with above or below median share of Catholics (share > 32 percent). We find that the fertility effects are more pronounced in counties where the majority voted for the Democrats in the 1920 or the 1932 presidential elections. In 1920, support for the Democrats was fully concentrated in the Southern states, whereas in 1932 the Democrats had the majority in all states. In the eleven Southern states[18], the estimated effect is almost twice as large as in the Northern states (also relative to the respective group means). Moreover, we do not find any systematic difference in the effect of exposure on fertility between counties facing a higher or lower local intensity of the Great Depression. The point coefficients are similar for counties with above and below median retail sales growth from 1929-1939; they are somewhat larger in areas with below median increase in social spending but, at the same time, somewhat smaller in areas with above median unemployment in 1930. Finally, the effects are estimated to be somewhat smaller in areas with an above median share of veterans in the 1930 population (see Table A.6 in the Appendix).

5. Evidence from vital statistics

Although the Census data provide a good starting point, there are several limitations that must be taken into account. First, we are unable to directly observe yearly births and instead rely on the number of a woman's own children below the age of five in the household. Secondly, the Census data are only collected every decade, leaving us with information from only three points in time: 1920, 1930, and 1940. Thirdly, we are unable to observe deaths although fetal deaths, infant deaths and maternal deaths are interesting outcome variables in our setting. To overcome these limitations, we complement the empirical analysis with a second approach in which we use yearly vital statistics on the universe of live births, stillbirths and infant deaths at the county level as well as causes of death data at the city level.

We create a balanced panel of counties over the period from 1925 to 1939. As explained in Section 3, the number of counties with complete vital statistics increases over the years. Therefore, the later we start the balanced panel, the more counties we can draw on. The earlier we start the balanced panel, the more years we can draw on. The restriction to a 15 years balanced panel from 1925 to 1939 maximizes the total number of observations; it leaves us with 27,360 observations from 1,824 counties.[19] To make treatment and control

[18]These are South Carolina, Mississippi, Georgia, Louisiana, Florida, Alabama, Texas, Virginia, Arkansas, North Carolina and Tennessee.

[19]In line with Sun and Abraham (2021), we drop counties treated already prior to the first year of observation.

counties more similar, we perform a matching on observables approach in the baseline period 1925.[20]. In particular, we run a probit model in which we regress a dichotomous variable which indicates whether a county gets treated in our period of observation on all variables that we also used as predictor variables in the roll-out regressions of Appendix Table A.1. Additionally, we include the stillbirth rate, the infant mortality rate as well as the all-age mortality rate (all observed in 1925) as covariates. We then match untreated counties to treated counties using nearest neighbor propensity score matching techniques and drop unmatched counties from the data set. Proceeding like this, we hold the composition of the sample constant over calendar time and ensure that treatment and control counties are comparable in observed baseline characteristics.[21] We checked that the Census results using Equation 1 are fully confirmed if we restrict the Census sample to women living in counties that are included in the balanced county panel or the matched balanced county panel (see Appendix Table A.3).

A natural consequence of our decision to keep the sample balanced in calendar time is that the sample is unbalanced in event time. Figure 8 shows how many treated counties are observed in each relative time period. The number of treated counties slightly increases from 800 in relative time −6 to 888 in relative time 0 and then decreases to 613 in relative time +3, 307 in relative time +6 and 73 in relative time +10. As a complementary statistic, Figure A.5 in the Appendix shows the shares of cohorts contributing to each relative time, where a cohort is defined as the group of counties who established a birth control clinic in the same year. We have decided to balance in calendar time and not in event time because this allows us to choose the baseline year 1925 for the matching of treated and control counties. Using a sample that is balanced in event time would need additional choices and assumptions on the year we should use for the matching of control counties. Moreover, while we could in principle balance the treated counties in event-time, the control counties would not be balanced since not all counties are observed in the historical data over the entire period of observation as we show in Figure 5.

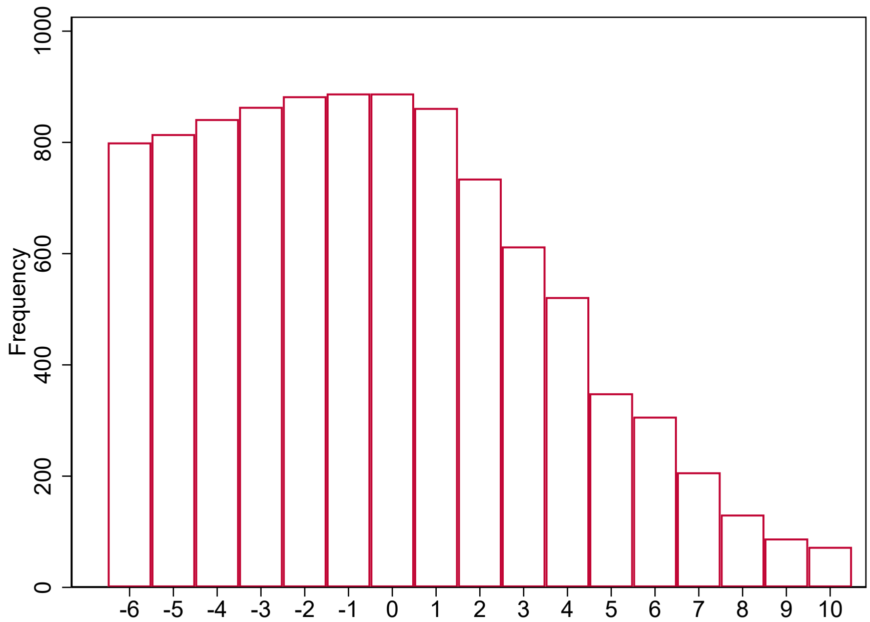
5.1. Identification strategy

With yearly county-level data, our setting boils down to a typical "staggered roll-out design", in which a binary absorbing treatment (i.e., a birth control clinic) is introduced

[20]We use a matching approach to restrict the sample to comparable units as we cannot control for time-varying covariates because detailed county level covariates are not available on an annual basis in our period of observation. Moreover, Sant'Anna and Zhao (2020) show that the inclusion of time-varying covariates in two-way fixed effects designs comes with several stringent assumptions in addition to the standard parallel trends and no anticipation assumptions, e.g, treatment effect homogeneity and parallel trends in each of the included covariates.

[21]Appendix Table A.7 presents the results of the probit model underlying the matching approach and compares treatment and control counties in the total sample as well as treatment and control counties in the matched sample. Appendix Table A.8 compares the means of the fertlity and mortality variables in the matched balanced panel 1925-1939 to those of the unbalanced panel 1920 to 1939.

FIGURE 8 — Number of observed treated counties by event-time



Notes: Data source: U.S. Vital Statistics. The figure shows the number of treated counties by event-time where the event is the establishment of a birth control clinic in a county or an adjacent county. The sample is restricted to a matched balanced panel of counties for the years 1925 to 1939.

in different counties at different points in time. The standard two-way fixed effects model yields biased estimates in such a setting if treatment effects are heterogeneous across groups or over time (Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2020). In particular, in an event-study design, the estimated lead and lag coefficients are combinations of differences in trends from their own relative periods, from relative periods belonging to other periods included in the specification, and from other relative periods excluded from the specification (Sun and Abraham, 2021).[22]

We avoid these pitfalls by estimating an event-study applying Sun and Abraham's (2021) interaction weighted estimator. The estimator is constructed in three steps. First, we run a linear two-way fixed effects model where the outcome variable y_{it} is the birth rate, the stillbirth rate, or the infant mortality rate of county i in year t. The model includes county (α_i) and year fixed-effects (σ_t) as well as interactions D^l_{it} for relative event period l with cohort indicators e. A cohort is defined as the group of counties i whose treatment starts in the same year. The start of the absorbing treatment is determined by the year in which the first birth control clinic in a county or any adjacent county was established. We drop all observations from periods more than six years prior to and more than ten years after the event. Then, we follow Baker *et al.* (2022), omit the most negative relative period prior to the event ($l = -6$) and the year of treatment ($l = 0$) and estimate separate

[22]Figure A.6 in the Appendix shows with which weights the event times enter in the estimation of all lead and lag coefficients in a standard TWFE approach applied to our setting.

lead and lag coefficients for all remaining relative event periods. Cohorts that are always treated are excluded from the estimation, while the *never-treated* counties form the control group C. The error term ζ_{it} is clustered at the county level.

$$y_{it} = \alpha_i + \sigma_t + \sum_{e \notin C} \sum_{l \neq 0} \phi_{e,l}(\mathbf{1}\{E_i = e\} \cdot D^l_{it}) + \zeta_{it} \tag{3}$$

Second, we estimate the weights by sample shares of each cohort in the relevant event time periods l as explained in the following expression, where $T + 1$ is the total number of calendar period observations $t \in 0, ..., T$ per unit:

$$Pr\{E_i = e | E_i \in [-l, T - l]\} \tag{4}$$

Finally, we weigh the group-specific estimates $\hat{\phi}_{e,l}$ from Equation 3 with the estimated weights from Equation 4 to obtain the interaction weighted estimator $\hat{v}_g$:

$$\hat{v}_g = \frac{1}{|g|} \sum_{l \in g} \sum_{e} \hat{\phi}_{e,l} \hat{Pr}\{E_i = e | E_i \in [-l, T - l]\} \tag{5}$$

This estimator has a clear interpretation since the weights sum to one for each relative time and are non-negative. The interaction weighted estimator depicts the average causal effect of birth control clinics on the treated counties under the assumptions of parallel trends in absence of the treatment and no anticipation of treatment.

We provide evidence for the validity of the parallel trends assumption by inspecting pre-trends. In particular, if the lead coefficients in the event-study are estimated to be zero, this makes it plausible that trends would also have been parallel in the periods after the opening of a birth control clinic if the birth control clinic had never opened. To investigate the validity of the no anticipation of treatment assumption, we inspect whether there are any conspicuous changes shortly before the opening of the first birth control clinic.

5.2. Effects on fertility

Figure 9 shows the event study estimates for the impact of birth control clinics on the crude birth rate, i.e., the number of births per 1,000 population. The small pre-treatment coefficients suggest that treatment group counties follow a similar birth rate trend as never-treated counties, which corroborates the validity of the parallel trends assumption. Only after the opening of a birth control clinic, we find a conspicuous and highly significant drop of the birth rate in treatment group counties as compared to control group counties. This negative effect is persistent and grows in the course of ten years after the opening of the birth control clinic.[23] Averaged over the period of ten years, birth control clinics reduce

[23]This increase of treatment effects with event time can partly be explained by the opening of further clinics in the county (see Figure A.2 in the Appendix).

the local annual birth rate by 0.628 births, which amounts to 3.4 percent as compared to the year the birth control clinic was established in a county or an adjacent county.[24]

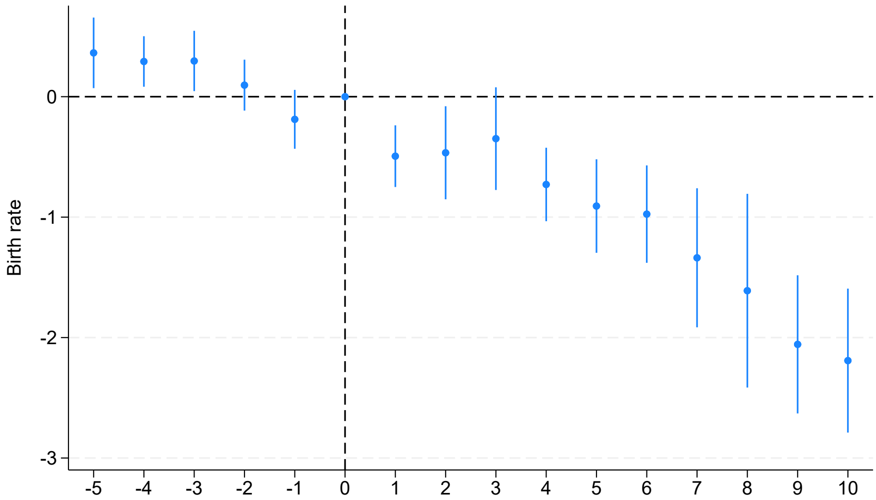
Comparing these estimates to the Census data estimates, we find that the estimated effect sizes are indeed similar. While the Census data suggest that exposure to birth control clinics accounts for 5.4 percent of the decline of the number of a woman's children below the age of five across the U.S. in the observed period, the vitality statistics suggest that birth control clinics account for 6.1 percent of the decline in the crude birth rate.[25] Although the estimated effect sizes seem to be not too distinct from each other, note that comparability might suffer from the fact that both the time period and the set of counties slightly differs between the Census data and the vital statistics estimations for data availability reasons. While we draw on data from all counties in the period from 1920 to 1940 in the Census data, the vital statistics restrict us to a panel of counties that consistently report vital statistics from 1925 to 1939, from which we additionally drop unmatched counties. Moreover, whereas the analysis based on the Census data uses the number of a married woman's surviving children below the age of five in the household, the vital statistics use live births per 1,000 population as the fertility measure.

How large are these effects as compared to the effects of the modern birth control pill? Looking at the advent of the modern contraceptive pill, Bailey (2010) finds that the Pill explains about 40 percent of the total decline of the marital fertility rate from 1955 to 1965. As compared to the estimated effects of the Pill, the estimated effects of the early birth control clinics seem to be small. This makes sense since the Pill is even more effective than a diaphragm, and more women could get access to the Pill than could get access to one of the 639 birth control clinics established across the U.S. At the same time, however, note that once a woman got a fitted diaphragm and was instructed how to use it, she could have used it for years with an effectiveness reaching 90 to 95 percent (Hajo, 2010; Lane *et al.*, 1976). Similar to the Pill, the woman was in control of the contraceptive device and thus did not have to rely on man's cooperation. Also remember that birth control clinics offered their services irrespective of the women's ability to pay. Thus, budget constraints did not prevent women from obtaining diaphragms in birth control clinics while the Pill was prohibitively expensive limiting its use at least in poorer strata of the population. U.S. federally funded family planning programs tried to tackle this problem in the late 1960s. And indeed, Bailey (2012) shows that from 1964 to 1973, these programs

[24]Note that the weights by which event time coefficients 1 to 10 enter this average depend on the number of observed treated counties at each event time as depicted in Figure 8. If we instead give each event time coefficient the same weight, the respective effect would be 1.112, or 6.1 percent.

[25]In the 1925–1939 balanced panel, the total decline in the birth rate amounts to 4.04 births per 1,000 population. For the 49.7 percent of the ever-treated counties, the average number of years since the establishment of the first birth control clinic by the end of our period of observation is 4.45 years. Over the first five years after the establishment of a birth control clinic, we estimate an average fertility effect of -0.497. Thus, birth control clinics can explain (0.497*0.497)/4.04=6.1 percent of the total decline of the crude birth rate.

Figure 9 — The impact of birth control clinics on fertility (Sun and Abraham, 2021)



Notes: Data source: U.S. Vital Statistics. The figure shows the dynamic impact of the establishment of the first birth control clinic in a county (or an adjacent county) on the birth rate (births per 1,000 population). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). The control group consists of never-treated counties. Standard errors are clustered at the county level. The whiskers mark the 95 percent confidence band. The sample is restricted to a matched balanced panel of counties for the years 1925 to 1939.

increased the use of the Pill by 5 percentage points and reduced fertility of women below 150 percent of the poverty line by 19 to 30 percent. Finally, whereas birth control clinics targeted mothers, the Pill was also available for women without any children. Overall, these cautious considerations suggest that the birth control clinic effects are considerably smaller than the effects of the modern contraceptive pill, yet still economically meaningful and not of an unreasonable size.

5.3. Effects on stillbirths and infant deaths

If birth control clinics reduce fertility by increasing birth spacing, this might result in a lower incidence of stillbirths and infant deaths. The link between birth spacing and perinatal outcomes has been persuasively documented using both historical and recent data. Using the completed reproductive histories for village populations in the 18th and 19th century, Knodel and Hermalin (1984) provide evidence for an adverse impact of short interpregnancy intervals on infant mortality. Molitoris (2017) confirms this finding in Swedish data from the early 20th century exploiting within-family variation in birth intervals. Reviewing research published in medical and (reproductive) health journals from 1966 to 2006, Conde-Agudelo *et al.* (2006) conclude that interpregnancy intervals shorter than 18 months or longer than 59 months are associated with a range of adverse perinatal outcomes. More recently, Gupta *et al.* (2019) show in a case control study that shorter intervals between pregnancies increase the risk of stillbirths. Conde-Agudelo *et al.*

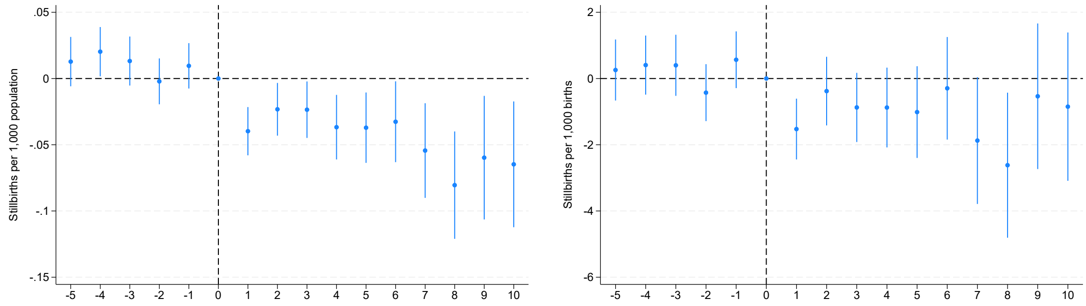
(2012) review causal biological mechanisms underlying these effects. They find that an inadequate time to recover from the insufficient repletion of maternal folate resources, vertical transmission of infections to the fetus, and transmission of infectious diseases among siblings seem to be important channels. Since pregnancy reduces milk flow, short interpregnancy intervals might also impede the survival chances of the previous infant in the absence of sterile and nutritionally adequate substitutes to breast milk (Knodel and Hermalin, 1984). Finally, birth control clinics might also have reduced stillbirths and infant deaths by reducing neonaticide and infanticide in desperate families; while we are not aware of encompassing data of infanticide in our period of observation, Oberman (2002) uses Chicago as a case study to argue that infanticide, and in particular neonaticide, were not uncommon in the late 19th and early 20th century.

Figure 10 depicts the impact of birth control clinics on stillbirths. In the left panel, we use the number of stillbirths per 1,000 population, while in the right panel, we use the number of stillbirths per 1,000 births (including stillbirths) as the outcome variable of the event-study. For both outcome variables, the pre-treatment coefficients fluctuate around zero and are mostly insignificant, which provides evidence for the validity of the key identifying assumptions. After the establishment of a birth control clinic, we observe a significant decrease of stillbirths. We observe this decrease in particular for the number of stillbirths per 1,000 population but also for the number of stillbirths per 1,000 births.[26]. The latter result suggests that the effect is not just a mechanical effect arising from a decrease in pregnancies; rather, it seems that birth control clinics avert particularly high risk pregnancies. Averaged over ten years, birth control clinics significantly reduce the number of stillbirths per 1,000 population by 0.035, and the number of stillbirths per 1,000 births by 1.028, which amounts to a decline of 5.6 percent and 3.2 percent, respectively, as compared to the year the birth control clinic was established in a county or an adjacent county.

In a next step, we investigate the impact of birth control clinics on infant mortality. As we can see in Figure 11, all five pre-treatment coefficients are insignificant and virtually zero, which again corroborates the validity of the common trends and the no-anticipation assumption. After the opening of a birth control clinic, the infant death rate drops; in the following years, the estimated negative effect of birth control clinics on infant mortality steadily increases. Since we define the infant death rates over the number of births, these estimates suggest that birth control clinics indeed improve the average health of the babies born. Again, this finding suggests that birth control clinics particularly avert the type of births that are characterized by increased health risks. Averaged over the period of ten years, birth control clinics significantly reduce the local infant death rate by 1.965 deaths, which amounts to 3.5 percent as compared to the year the birth control clinic was

[26]Even if not all single lag coefficients reach conventional significance levels, the compound post effect is significantly negative.

FIGURE 10 — The impact of birth control clinics on stillbirths (Sun and Abraham, 2021)



Notes: Data source: U.S. Vital Statistics. The figure shows the dynamic impact of the establishment of the first birth control clinic in a county (or an adjacent county) on stillbirths per 1,000 population (left panel) and stillbirths per 1,000 births incl. stillbirths (right panel). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). The control group consists of never-treated counties. Standard errors are clustered at the county level. The whiskers mark the 95 percent confidence band. The sample is restricted to a matched balanced panel of counties for the years 1925 to 1939.

established in a county or an adjacent county.[27] Taken together, birth control clinics do have significant and meaningful health effects; yet, at the same time, the estimated effects are definitely smaller than those of public health infrastructure such as sewerage systems and waterworks in the early 20th century. Alsan and Goldin (2019) show that sewerage and waterworks in combination reduced infant mortality by 48 percent in Massachusetts in the period of 1880 to 1920. Cutler and Miller (2005) find similarly large effects (43 percent) for water filtration in a sample of 13 large American cities in the period of 1900 to 1936, while Anderson *et al.* (2022) re-estimate the effects based on an extended sample of 25 large American cities and find considerably smaller effects of 11 percent.

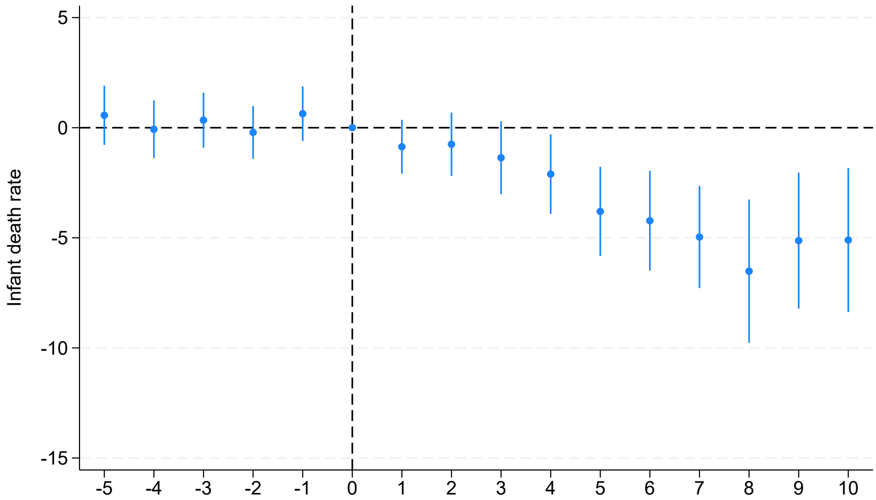
Since birth control clinics reduce infant mortality, the negative effect on fertility identified in the Census data might be a lower bound estimate. This is because the Census fertility measure, i.e., the number of a woman's own children below five in the household, is driven up in treatment as compared to control households due to reduced infant mortality. Thus, the isolated negative fertility effect in the vital statistics should be larger than the one estimated in the Census data, which is in line with our findings.

5.4. Effects on maternal mortality

As birth control clinics increase birth spacing and reduce the number of births, they might also reduce maternal deaths. In their review article, Conde-Agudelo *et al.* (2012) report that short interpregnancy intervals might result in maternal nutritional depletion with negative health consequences for the mother; yet, they also clarify that the empirical evidence for this channel is still inconclusive. Moreover, access to contraceptives via birth

[27] Complementing our historical evidence, Flynn (2024) finds that a program that expanded access to long-acting reversible contraceptives to lower income women reduced infant mortality in Colorado during the period 2009-2015.

Figure 11 — The impact of birth control clinics on infant mortality (Sun and Abraham, 2021)



Notes: Data source: U.S. Vital Statistics. The figure shows the dynamic impact of the establishment of the first birth control clinic in a county (or an adjacent county) on the infant death rate (infant deaths per 1,000 births). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). The control group consists of never-treated counties. Standard errors are clustered at the county level. The whiskers mark the 95 percent confidence band. The sample is restricted to a matched balanced panel of counties for the years 1925 to 1939.

control clinics might have reduced unprofessional abortions and thereby reduced health risks. Finally, maternal mortality might also drop simply because the number of births is reduced.

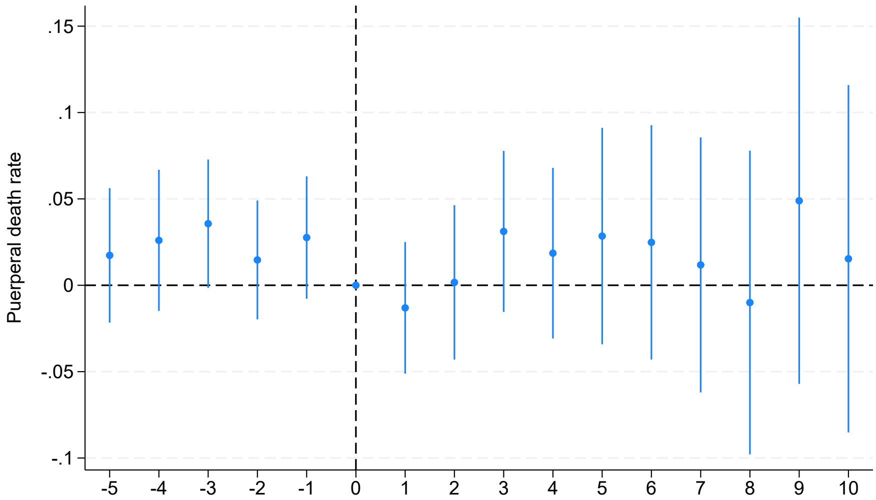
Figure 12 depicts the effect of birth control clinics on the puerperal death rate, i.e., the number of puerperal deaths per 1,000 women aged 15 to 49, at the level of 316 cities using the interaction weighted estimator by Sun and Abraham (2021). The small and insignificant pre-treatment coefficients provide evidence for the validity of the common trends assumption. At the same time, the post post-treatment coefficients hover around zero and non of them reaches conventional significance levels. Thus, we do not find any evidence for negative effects of birth control clinics on puerperal deaths.

5.5. Robustness tests

As a robustness test, we check whether our findings are confirmed if we use a 20 years balanced panel instead of the 15 years balanced panel of counties. Overall, the 20 years balanced panel leaves us with 25,900 observations from 1,295 counties. Again, we apply nearest neighbor matching and drop counties that are unmatched, which leaves us with 18,840 observations from 924 counties.[28] Figure A.7 in the Appendix reports the event study estimates on this alternative sample and fully confirms all previous findings. While

[28] Table A.8 in the Appendix shows that the fertility and mortality rates in this sample are comparable to those of the 15 years balanced panel of the main specification.

FIGURE 12 — The impact of birth control clinics on puerperal deaths (Sun and Abraham, 2021)



Notes: Data source: City-level causes of death data. The figure shows the dynamic impact of the establishment of the first birth control clinic in a county (or an adjacent county) on the puerperal death rate (puerperal deaths per 1,000 female population aged 15 to 49). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). The control group consists of never-treated cities. Standard errors are clustered at the county level. The whiskers mark the 95 percent confidence band. The sample is restricted to a balanced panel of cities that consistently report causes of deaths from 1920 to 1937.

the effects of birth control clinics on the crude birth rate are very similar, the effects on stillbirths and infant mortality are even more pronounced than in the 15 years balanced panel.

In an alternative specification, we use the last treated cohort instead of matched never-treated counties as the control group. Since the number of counties with a birth control clinic opening in 1939 is small, the resulting estimates might be affected by outliers. Therefore, we extend the last treated cohort to cohorts that opened a birth control clinic either in 1938 or in 1939 and drop these two years from the sample in line with Sun and Abraham (2021).[29] Figure A.8 in the Appendix presents the results of this alternative specification. The estimated effect of birth control clinics on the birth rate is somewhat attenuated; it becomes significant only seven years after the opening of the birth control clinic. The effects of birth control clinics on stillbirths and the infant mortality rate are consistently negative but less precisely estimated than in the main specification, likely due to the lower number of observations.

Finally, we check whether the results are confirmed using the estimator proposed by Callaway and Sant'Anna (2021) instead of Sun and Abraham's (2021) interaction weighted estimator. As in the main specification, we use a 15 years balanced panel from 1925 to 1939 and restrict the control group to never-treated counties. Moreover, we again use

[29]Table A.8 in the Appendix shows that also this sample hardly differs from the 15 years balanced panel of the main specification.

nearest neighbor propensity score matching techniques and drop unmatched counties from the data set. All findings are confirmed using this alternative estimator (see Appendix Figure A.9).

5.6. Validity checks and discussion

We use the county level all-age mortality rates and the city level typhoid fever death rates to rule out that any estimated negative mortality effects on unborn children and infants are confounded by general improvements of the public health infrastructure. In particular, the roll-out of sewerage systems and waterworks was among the most important health infrastructure measures at the end of the 19th and beginning of the 20th century (Alsan and Goldin, 2019; Costa, 2015). Whereas Cutler and Miller (2005) argue that clean water technologies substantially reduced total mortality in U.S. cities in the early 20th century, Anderson *et al.* (2022) re-evaluate the impact and find smaller effects that are concentrated in typhoid fever mortality and infant mortality. Therefore, if the establishment of birth control clinics is confounded by the roll-out of sewerage systems and waterworks, we might find negative associations with the all-age mortality rate and we should definitely find negative associations with the waterborne typhoid fever mortality rate. To check whether this is the case, we use the all-age mortality rate (without infant deaths) and the typhoid mortality rate as the outcome variables of our event-study specification. We find that for both outcomes, all pre-treatment coefficients are small and insignificant (see Appendix Figure A.10). After the opening of a birth control clinic, we do not find any consistent evidence for significant negative associations with the all-age mortality rate. Indeed, the coefficients hover around zero and are insignificant in eight out of ten post periods; if we combine all lag coefficients into a single post indicator, the estimated post effect is insignificant (p-value 0.989). For typhoid mortality, none of the ten post coefficients is significant while the point estimates are even positive. Thus, these findings provide evidence that a general improvement in counties' health infrastructure and in particular the provision of sewerage systems and waterworks does not confound the negative impact of birth control clinics on stillbirths and infant mortality.

It is also unlikely that the Sheppard Towner Act, which went into effect in late 1921 and expired in 1929, confounds the impact of the birth control clinics on mortality. Lemons (1969) reports that the Act resulted in 3,000 child and maternal health care centers mostly in small cities and rural areas. Moehling and Thomasson (2012) provide state level evidence that the Sheppard Towner Act might have reduced infant mortality. However, two points speak against the Sheppard Towner Act being an important confounder of the birth control clinics effects. First, our stillbirths effects hold in a subsample of urban counties which are less targeted by the Sheppard Towner Act; only the effect on the infant death rate remains insignificant with a p-value of 0.122 (see Appendix Table A.9).[30] Second, while the Sheppard Towner Act was repealed in 1927 and fully expired in 1929, our fertility and mortality effects are identified in a balanced sample of counties in the period 1925 to 1940, and 92 percent of all counties that established birth control clinics in this sample did so in the 1930s. Therefore, the years in which the Sheppard Towner Act was still in effect mostly coincide with pre-treatment years. Almost all estimates show zero pre-treatment correlations of birth control clinics with fertility and mortality, while the effects materialize shortly after the birth control clinics were established. Therefore, we conclude that the Sheppard Towner Act is unlikely to explain the pattern of results we find.

Another concern is that welfare programs for poor families confound the estimated birth control clinic effects on mortality. Mothers' Pensions introduced in 1911 were the United States' first welfare program of this kind. The program targeted poor families who suffered from the loss or disability of the breadwinner. Aizer *et al.* (2016) report that benefits amounted to 12 to 25 percent of family income with a median benefit duration of three years. Guidelines on eligibility and funding considerably differed across states and it was up to counties to implement the benefits. Many counties did not implement Mothers' Pensions despite state laws. If counties provided Mothers' Pensions, families were severely underserved with two thirds of targeted families not receiving any benefits (Lundberg, 1926). Most importantly, Aizer *et al.* (2016) show that only 2 percent of the children whose mothers did successfully apply for mothers' pensions were less than one year old. Since we focus on infant mortality, i.e., mortality of children below the age of one, this makes it highly unlikely that our mortality effects are confounded by Mothers' Pensions. After the Social Security Act of 1935, Aid to Dependent Children (ADC) replaced Mothers' Pensions. Moreover, New Deal relief programs supported poor families during the Great Depression (Fishback *et al.*, 2007). To empirically check that the expansion of social welfare programs in the 1930s does not confound the birth control clinic estimates, we use data from Fishback *et al.* (2005) on average per capita public works and relief spending in the periods 1933-1935 and 1933-1939. We split the sample into counties with above and below median per capita public works and relief spending in the period 1933-1939. As an alternative, we divide this variable by per capita public works and relief spending in the period 1933-1935 to capture social spending growth after 1935, and split the sample at the median of social spending growth. Then, we re-run our event study analyses separately on the sample of counties with a comparatively large and with a comparatively small prevalence of welfare benefits. Again, we find negative birth control clinic effects on fertility and stillbirths in both groups of counties; the infant mortality effects are consistently negative in both groups but do not reach conventional

[30] Since big city counties represent only 3 percent of all counties in our matched sample, we cannot restrict the sample to this small group of counties using the county level vital statistics.

significance levels in counties with a smaller prevalence of welfare benefits (see Appendix Table A.9). Thus, we conclude that early social welfare programs are unlikely to confound our effects.

Finally, one might be concerned that the local intensity of the Great Depression confounds the estimates. While Dehejia and Lleras-Muney (2004) show that babies born in economic downturns of the 1970s to 1990s had better health outcomes, partly due to selection, partly due to improved health behaviors during recessions, Fishback *et al.* (2007) find that many causes of death were indeed pro-cyclical also in the 1930s; but this was not the case for infant deaths. We use data on retail sales from Fishback *et al.* (2005) and the unemployment rate in 1930 as proxies for the local intensity of the Great Depression. To improve the comparability of counties, we then create a subsample of counties heavily affected by the Great Depression (below median retail sales growth, or above median unemployment in 1930) and a subsample of counties moderately affected by the Great Depression (above median retail sales growth, or below median unemployment in 1930). It turns out that the birth control clinic effects on fertility and mortality show up both in counties heavily and in counties moderately affected by the Great Depression (see Appendix Table A.9). We take this result as evidence that the local intensity of the Great Depression does not confound the estimated birth control clinic effect.

6. Conclusions

This study contributes new insights into the factors that drove the U.S. demographic transition at the beginning of the 20th century and sheds light on the impact of modern contraception, especially professionally fitted diaphragms, which had the advantage of not requiring men's cooperation, before the introduction of the birth control pill in 1960. While there is a rich historical and sociological literature on Margaret Sanger's life and her movement, this work is the first quantitative assessment of the first family planning initiative in the U.S. Our research draws on a unique combination of newly digitized data on the roll-out of Sanger's birth control clinics, full-count Census data, as well as county and city-level vital statistics. To assess the causal impact of birth control clinics on fertility, stillbirths, infant mortality, and maternal deaths, we employ event-study methods that leverage the staggered roll-out of clinics across U.S. counties.

Historical Census data estimates indicate that exposure to a birth control clinic of about six years which corresponds to the average exposure time of women at the end of your observation period reduces the number of a woman's children below the age of five in the household by 3.7 percent. Event-study estimates based on county level vital statistics confirm the negative impact of birth control clinics on fertility. Birth control clinics reduce the local crude birth rate by 3.2 percent within ten years after their establishment. Overall, birth control clinics account for 5.4–6.1 percent of the total decline in the crude birth rate across the U.S. between 1920 and 1940. In addition, we find evidence that birth control clinics had a significant and meaningful negative effect on the local incidence of stillbirths per 1,000 population (-5.6 percent), stillbirths per 1,000 births (-3.2 percent), and infant mortality (-3.5 percent) in a period of ten years after establishment. These findings are in line with the notion that birth control clinics increase birth spacing and thus particularly avert births that pose considerable health risks. We do not find any evidence for negative effects of birth control clinics on maternal deaths, particularly puerperal deaths. Parallel pre-treatment trends provide evidence for the validity of the key identifying assumptions. Various additional validity and specification checks including placebo outcomes, alternative control groups, and subsample analyses confirm the findings.

Overall, the findings suggest that the relaxation of supply side policies by birth control clinics reduced constraints on the demand side, i.e., in a Beckerian framework (Becker, 1981) birth control clinics reduced the cost of increasing interpregnancy intervals. In particular, they empowered women to take control over their fertility, whereas other methods popular at that time, such as condoms, withdrawal or abstinence required the cooperation of men (and were less effective). Our results demonstrate the power of family planning in a context where parents, and especially mothers desire smaller family size and longer birth intervals but are constrained to implement these preferences. The study also indicates that family planning can have substantial indirect effects on health. These findings are likely applicable to much of the developing world, in contexts where women desire smaller families but give birth to many children at a young age in a relatively short period.

References

Aaronson, D., Lange, F. and Mazumder, B. (2014). Fertility transitions along the extensive and intensive margins. *American Economic Review*, **104** (11), 3701–3724.

Ager, P., Feigenbaum, J. J., Hansen, C. W. and Tany, H. R. (2024). How the other half died: Immigration and mortality in US cities. *The Review of Economic Studies*, **91** (1), 1–44.

—, Herz, B. and Brueckner, M. (2020). Structural change and the fertility transition. *The Review of Economics and Statistics*, **102** (4), 806–822.

Aizer, A., Eli, S., Ferrie, J. and Lleras-Muney, A. (2016). The long-run impact of cash transfers to poor families. *American Economic Review*, **106** (4), 935–971.

Alsan, M. and Goldin, C. (2019). Watersheds in child mortality: The role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy*, **127** (2), 586–638.

Anderson, D. M., Kofi Charles, K. and Rees, D. J. (2022). Reaximining the contribution of public health efforts to the decline in urban mortality: Reply. *American Economic Journal: Applied Economics*, **14** (2), 166–169.

Bailey, M. J. (2006). More power to the pill: The impact of contraceptive freedom on women's life cycle labor supply. *The Quarterly Journal of Economics*, **121** (1), 289–320.

— (2010). Momma's got the pill: How Anthony Comstock and Griswold v. Connecticut shaped US childbearing. *American Economic Review*, **100** (1), 98–129.

— (2012). Reexamining the impact of family planning programs on US fertility: Evidence from the war on poverty and the early years of Title X. *American Economic Journal: Applied Economics*, **4** (2), 62–97.

— (2013). Fifty years of US family planning: New evidence on the long-run effects of increasing access to contraception. *Brookings Papers on Economic Activity*, **Spring**, 341–409.

—, Clay, K., Fishback, P., Haines, M., Kantor, S., Severnini, E. and Wentz, A. (2016). *U.S. County-Level Natality and Mortality Data, 1915–2007*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

— and Collins, W. J. (2011). Did improvements in household technology cause the baby boom? evidence from electrification, appliance diffusion, and the Amish. *American Economic Journal: Macroeconomics*, **3** (2), 189–217.

— and Hershbein, B. J. (2018). US fertility rates and childbearing in American economic history, 1800-2010. In L. P. Cain, P. V. Fishback and P. W. Rhode (eds.), *The Oxford Handbook of American Economic History*, Oxford University Press.

—, Wanner Lang, V., Prettyman, A., Vrioni, I., Bart, L. J., Eisenberg, D., Fomby, P., Barber, J. and Dalton, V. (2023). *How Costs Limit Contraceptive Use among Low-Income Women in the in the U.S.: A Randomized Control Trial*. NBER Working Paper 31397, National Bureau of Economic Research, Cambridge, MA.

Baker, A. C., Larcker, D. F. and Wang, C. C. (2022). How much should we trust staggered differences-in-differences estimates? *Journal of Financial Economics*, **144** (2), 370–395.

Beach, B. and Hanlon, W. (2023). Culture and the historical fertility transition. *Review of Economics Studies*, **90** (4), 1669–1700.

Becker, G. S. (1981). *A Treatise on the Family*. Cambridge: Harvard University Press.

Bellou, A. and Cardia, E. (2021). Great depression and the rise of female employment: A new hypothesis. *Explorations in Economic History*, **80** (101383).

—, — and Lewis, J. (2024). From bust to boom: The great depression and women's fertility. *mimeo*.

Bleakley, H. and Lange, F. (2009). Chronic disease burden and the interaction of education, fertility, and growth. *The Review of Economics and Statistics*, **91** (1), 52–65.

Bourne Wahl, J. (1992). Trading quantity for quality: Explaining the decline in American fertility in the nineteenth century. In *Strategic Factors in Nineteenth Century American Economic History: A Volume to Honor Robert W. Fogel*, University of Chicago Press, pp. 375–397.

Callaway, B. and Sant'Anna, P. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, **225** (2), 200–230.

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

Chesler, E. (1992). *Woman of Valor: Margaret Sanger and the Birth Control Movement in America*. New York: Simon Schuster.

Conde-Agudelo, A., Rosas-Bermudez, A., Castaño, F. and Norton, M. H. (2012). Effects of birth spacing on maternal, perinatal, infant, and child health: A systematic review of causal mechanisms. *Studies in Family Planning*, **43** (2), 93–114.

—, — and Kafury-Goeta, A. C. (2006). Birth spacing and risk of adverse perinatal outcomes: A meta-analysis. *Journal of the American Medical Association*, **295** (15), 1809–1823.

Costa, D. (2015). Health and the economy in the United States from 1750 to the present. *Journal of Economic Literature*, **53** (3), 503–570.

Curtis White, K. J. (2008). Population change and farm dependence: Temporal and spatial variation in the U.S. Great Plains, 1900–2000. *Demography*, **45** (2), 363–386.

Cutler, D., Deaton, A. and Lleras-Muney, A. (2006). The determinants of mortality. *The Journal of Economic Perspectives*, **20** (3), 97–120.

— and Miller, G. (2005). The role of public health improvements in health advances: The twentieth-century United States. *Demography*, **42** (1), 1–22.

Cutright, P. and Jaffe, F. S. (1977). *Impact of Family Planning Programs on Fertility: The U.S. Experience*. New York: Praeger.

David, P. A., Sanderson, W. C. and Haines, M. R. (1986). Rudimentary contraceptive methods and the American transition to marital fertility control, 1855-1915. In S. L. Engerman and R. E. Gallman (eds.), *Long-Term Factors in American Economic Growth*, University of Chicago Press, pp. 307–390.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, **110** (9), 2964–2996.

Dehejia, R. and Lleras-Muney, A. (2004). Booms, busts, and babys' health. *The Quarterly Journal of Economics*, **119** (3), 1091—1130.

Engelman, P. C. (2011). *A History of the Birth Control Movement in America*. Santa Barbara, CA: Praeger.

Feigenbaum, J. J., Muller, C. and Wrigley-Field, E. (2019). Regional and racial inequality in infectious disease mortality in u.s. cities, 1900-1948. *Demography*, **56**, 1371–1388.

Fishback, P., Haines, M. R. and Kantor, S. (2007). Births, deaths, and new deal relief during the great depression. *Review of Economics and Statistics*, **89** (1), 1–14.

—, Horrace, W. C. and Kantor, S. (2005). The impact of new deal expenditures on local economic activity: An examination of retail sales, 1929-1939. *Journal of Economic History*, **65**, 36–71.

Flynn, J. (2024). Can expanding contraceptive access reduce adverse infant health outcomes? *Journal of Human Resources*.

Goldin, C. and Katz, L. F. (2002). The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy*, **110** (4), 730–770.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, **225** (2), 254–277, themed Issue: Treatment Effect 1.

Gordon, L. (2002). *The Moral Property of Women: A History of Birth Control Politics in America*. Champaign: University of Illinois Press.

Greenwood, J. and Seshadri, A. (2002). The U.S. demographic transition. *American Economic Review*, **92** (2), 153–159.

Grimm, M. (2021). Rainfall risk, fertility and development: evidence from farm settlements during the American demographic transition. *Journal of Economic Geography*, **21** (4), 593–618.

Guinnane, T. W. (2011). The historical fertility transition: A guide for economists. *Journal of Economic Literature*, **49** (3), 589–614.

Gupta, P. M., Freedman, A. A., Kramer, M. R., Goldenberg, R. L., Willinger, M., Stoll, B. J., Silver, R. M., Dudley, D. J., Parker, C. B. and Hogue, C. J. (2019). Interpregnancy interval and risk of stillbirth: a population-based case control study. *Annals of Epidemiology*, **35**, 35–41.

Guttmacher, A. (1947). Conception control and the medical profession: The attitudes of 3,381 physicians toward contraception and contraceptives they prescribe. *Human Fertility*, **12**, 1–10.

Hacker, J. D. (2003). Rethinking the "early" decline of marital fertility in the United States. *Demography*, **40** (4), 605–620.

Haines, M. R. (2000). The white population of the United States, 1790–1920. In M. R. Haines and R. H. Steckel (eds.), *A Population History of North America*, Cambridge, MA: Cambridge University Press, pp. 305–371.

— (2006). Historical statistics of the united states: Millennial edition. In S. B. Carter (ed.), *Historical Statistics of the United States: Millennial Edition*, Cambridge University Press, pp. 1–391.

— and Hacker, D. J. (2006). *The Puzzle of the Antebellum Fertility Decline in the United States: New Evidence and Reconsideration*. NBER Working Paper 12571, National Bureau of Economic Research, Cambridge, MA.

Hajo, C. M. (2010). *Birth Control on Main Street. Organizing Clinics in the United States, 1916–1939*. Champaign: University of Illinois Press.

Hansen, C. W., Jensen, P. S. and Lønstrup, L. (2018). The fertility decline in the United States: Schooling and income. *Macroeconomic Dynamics*, **22** (6), 1584–1612.

Holz, R. (2012). *The Birth Control Clinic in a Marketplace World*. Rochester: University of Rochester Press.

Horan, P. M. and Hargis, P. G. (1995). *County Longitudinal Template, 1840–1990*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Jones, L. E. and Tertilt, M. (2006). *An Economic History of Fertility in the U.S.: 1826-1960*. NBER Working Paper 12796.

Kitchens, C. T. and Rodgers, L. P. (2023). The impact of the WWI agricultural boom and bust on female opportunity cost and fertility. *The Economic Journal*, **133** (656), 2978–3006.

Knodel, J. and Hermalin, A. I. (1984). Effects of birth rank, maternal age, birth interval, and sibship size on infant and child mortality: Evidence from 18th and 19th century reproductive histories. *American Journal of Public Health*, **74** (10), 1098–1106.

Knowlton, C. (1832). *Fruits of Philosophy, Or the Private Companion of Adult People*. Peter Pauper Press.

Lahey, J. N. (2014a). Birthing a nation: The effect of fertility control access on the nineteenth-century demographic transition. *The Journal of Economic History*, **74** (2), 482–508.

— (2014b). The effect of anti-abortion legislation on nineteenth century fertility. *Demography*, **51** (3), 939–948.

— (2022). *Effects of Restrictive Abortion Legislation on Cohort Mortality: Evidence from 19th Century Law Variation*. NBER Working Paper 30201.

Lane, M. E., Arceo, R. and Sobrero, A. J. (1976). Successful use of the diaphragm and jelly by a young population: Report of a clinical study. *Family Planning Perspectives*, **8** (2), 81–86.

Lemons, S. J. (1969). The Sheppard-Towner Act: Progressivism in the 1920s. *The Journal of American History*, **55** (4), 776–786.

Lundberg, E. O. (1926). *Public Aid to Mothers with Dependent Children: Extent and Fundamental Principles. Children's Bureau Report*. Washington, DC: Government Printing Office.

Manson, S., Schroeder, J., Van Riper, D., Kugler, T. and Ruggles, S. (2022). *IPUMS National Historical Geographic Information System: Version 17.0*. Minneapolis, MS: IPUMS.

McCann, C. R. (1994). *Birth Control Politics in the United States. 1916–1945*. Ithaca: Cornell University Press.

Mellor, J. M. (1998). The effect of family planning programs on the fertility of welfare recipients: Evidence from Medicaid claims. *The Journal of Human Resources*, **33** (4), 866–895.

Moehling, C. M. and Thomasson, M. A. (2012). *Saving Babies: The Contribution of Sheppard-Towner to the Decline in Infant Mortality in the 1920s*. NBER Working Paper 17996.

Molitoris, J. (2017). The effect of birth spacing on child mortality in Sweden, 1878–1926. *Population and Development Review*, **43** (1), 61–82.

Molyneaux, J. W. and Gertler, P. J. (2000). The impact of targeted family planning programs in Indonesia. *Population and Development Review*, **26**, 61–85.

Oberman, M. (2002). Understanding infanticide in context: Mothers who kill, 1870–1930 and today. *The Journal of Criminal Law and Criminology*, **92** (3/4), 707–738.

Owen, D. O. (1876). Moral physiology, or: A brief and plain treatise on the population question.

Reed, J. (1978). *The Birth Control Movement and American Society: From Private Vice to Public*. Princeton: Princeton University Press.

Robinson, E. (1934). *The Presidential Vote, 1896–1932*. Stanford: Stanford University Press.

Ruggles, S., Flood, S., Goeken, R., Schouweiler, M. and Sobek, M. (2022). *IPUMS USA: Version 12.0*. Minneapolis, MS: IPUMS.

Sant'Anna, P. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, **219** (1), 101–122.

Steckel, R. H. (1992). The fertility transition in the United States: Tests of alternative hypotheses. In *Strategic Factors in Nineteenth Century American Economic History: A Volume to Honor Robert W. Fogel*, University of Chicago Press, pp. 351–374.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, **225** (2), 175–199, themed Issue: Treatment Effect 1.

Tone, A. (2002). *Devices and Desires: A History of Contraceptives in America*. New York City: Hill and Wang.

Udry, J. R., Bauman, K. E. and Morris, N. M. (1976). The effect of subsidized family planning services on reproductive behavior in the United States, 1969-1974. *Demography*, **13** (4), 463–478.

Wanamaker, M. H. (2012). Industrialization and fertility in the nineteenth century: Evidence from South Carolina. *The Journal of Economic History*, **72** (1), 168–196.

Weisbord, R. G. (1973). Birth control and the Black American: A matter of genocide? *Demography*, **10** (4), 571–590.

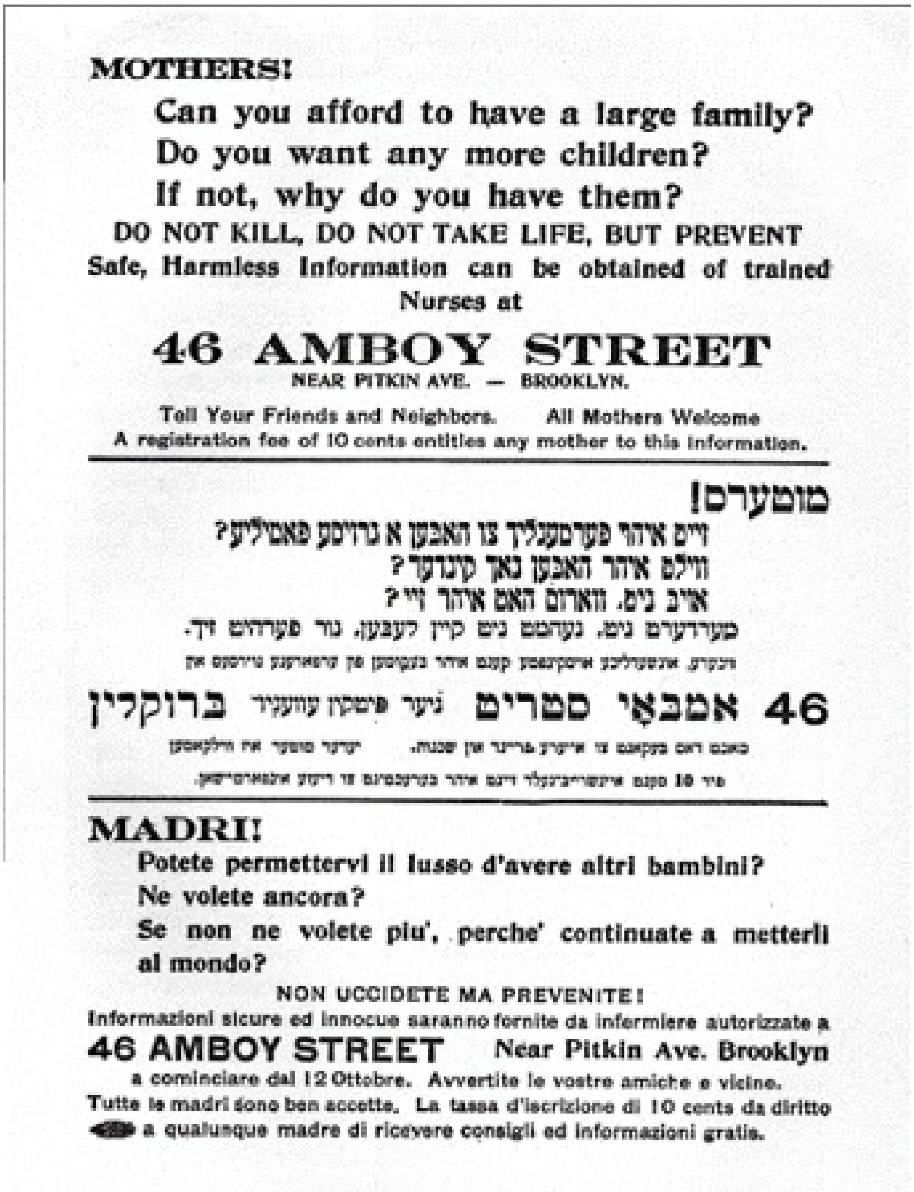
Wheeler, K. H. (2012). Infanticide in nineteenth-century Ohio. *Journal of Social History*, **31** (2), 407–418.

A. Appendix (For Online Publication)

This Online Appendix provides additional material discussed in the manuscript "The Impact of Margaret Sanger's Birth Control Clinics on Early 20th Century U.S. Fertility and Mortality" by Stefan Bauernschuster, Michael Grimm, and Cathy M. Hajo.

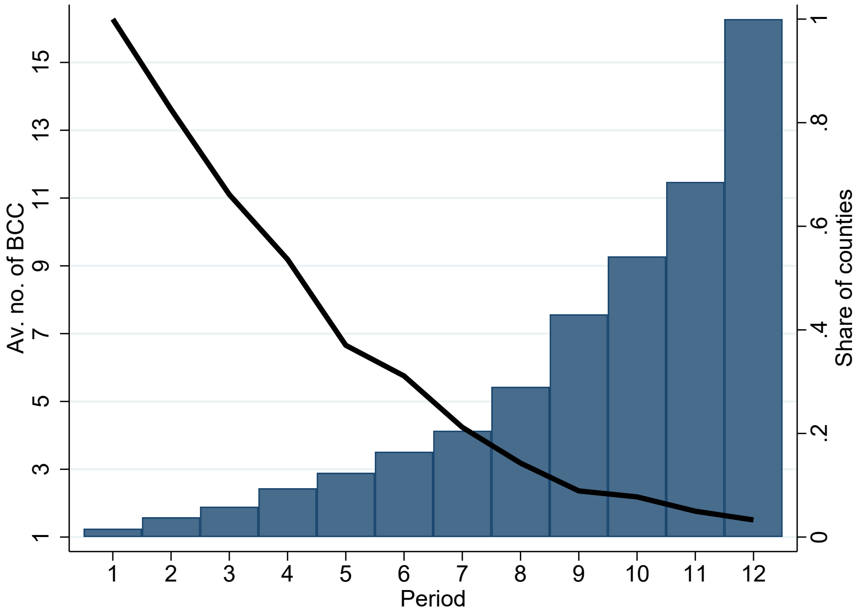
A.1. Additional Figures and Tables

FIGURE A.1 — Leaflet advertising the Sanger Clinic in 46 Amboy Street, Brooklyn



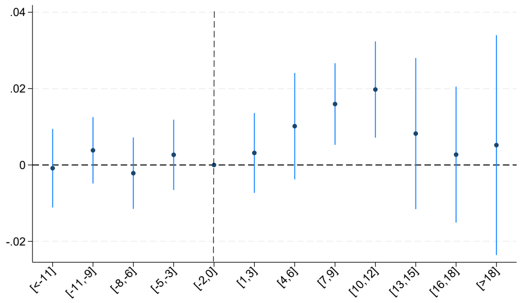
Notes: The photo shows a leaflet advertising the United States' first birth control clinic in English, Yiddish, and Italian.

FIGURE A.2 — Birth control clinics: dynamics of the roll-out



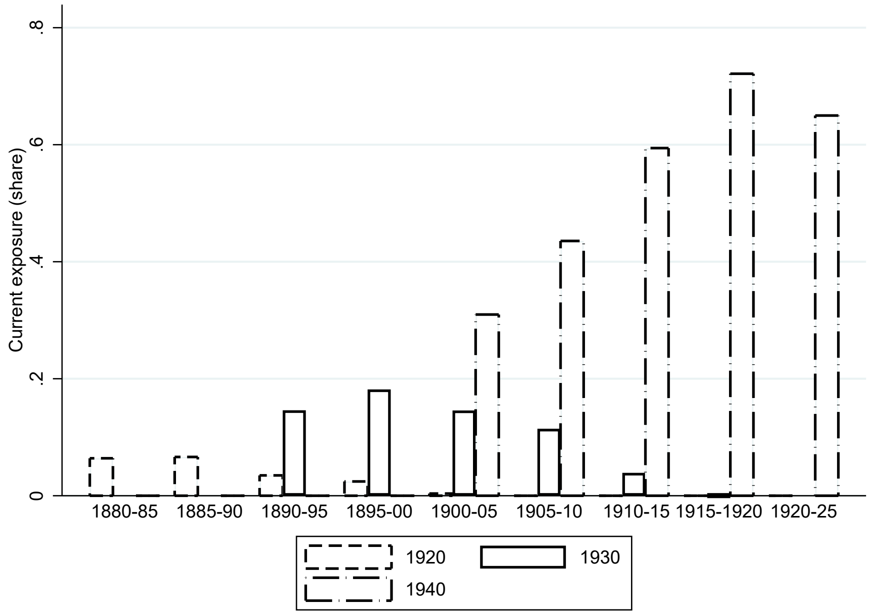
Notes: Data source: Birth control clinics statistics by Hajo (2010). The bars show the average number of birth control clinics in the own or adjacent county over the period since the first clinic was established. The black line shows the share of counties that experience the period depicted on the horizontal axis with at least one birth control clinic in the own or adjacent county. For example, in the fifth period after the establishment of the first birth control clinic, counties have on average nearly three clinics, but less than 30 percent of the counties experience five periods and more with a birth control clinic in our period of observation until 1940.

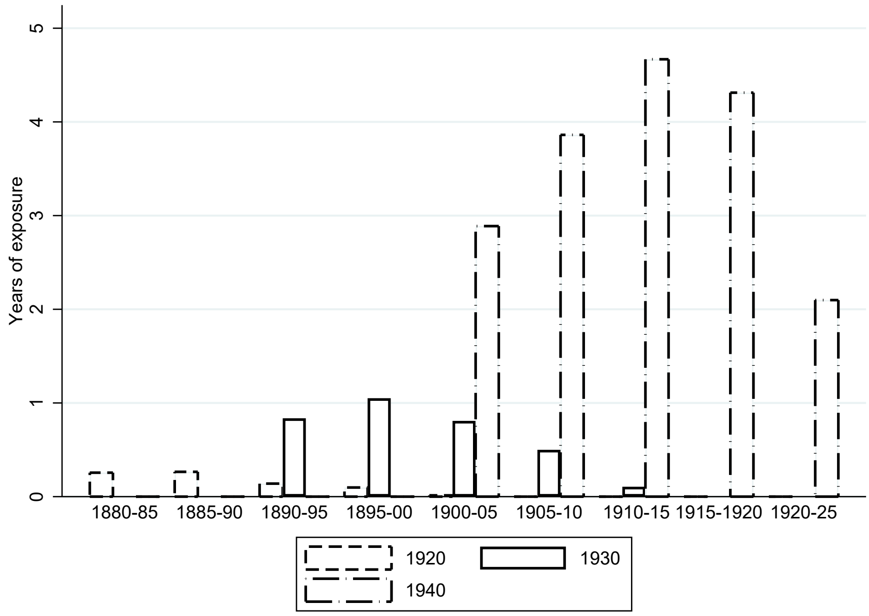
Figure A.3 — The impact of birth control clinics on female labor force participation



Notes: Data source: IPUMS US Census, 1920, 1930 and 1940. The figure shows event study estimates of the exposure to a birth control clinic on women's labor force participation. Treatment effects are estimated along the lines of equation 2. The whiskers mark the 95 percent confidence band.

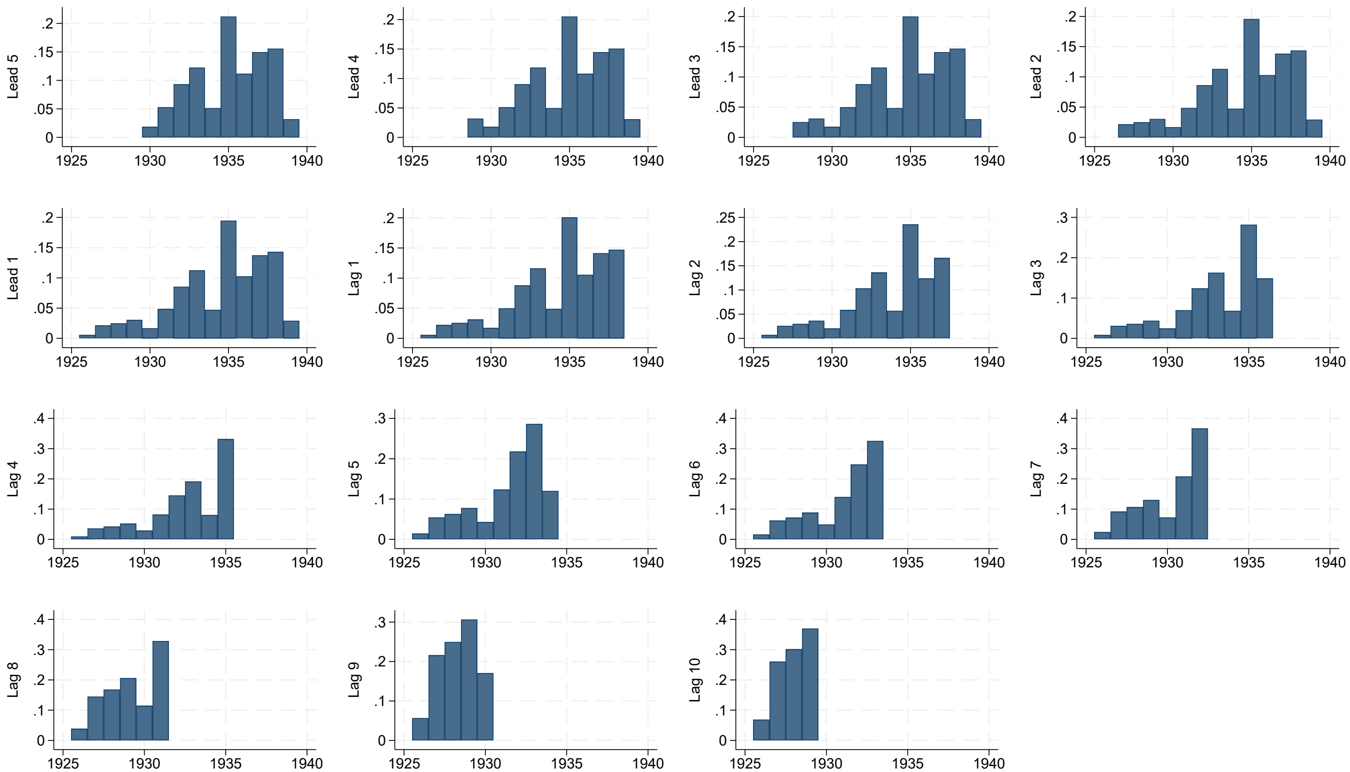
FIGURE A.4 — Exposure to a birth control clinic by cohort and year





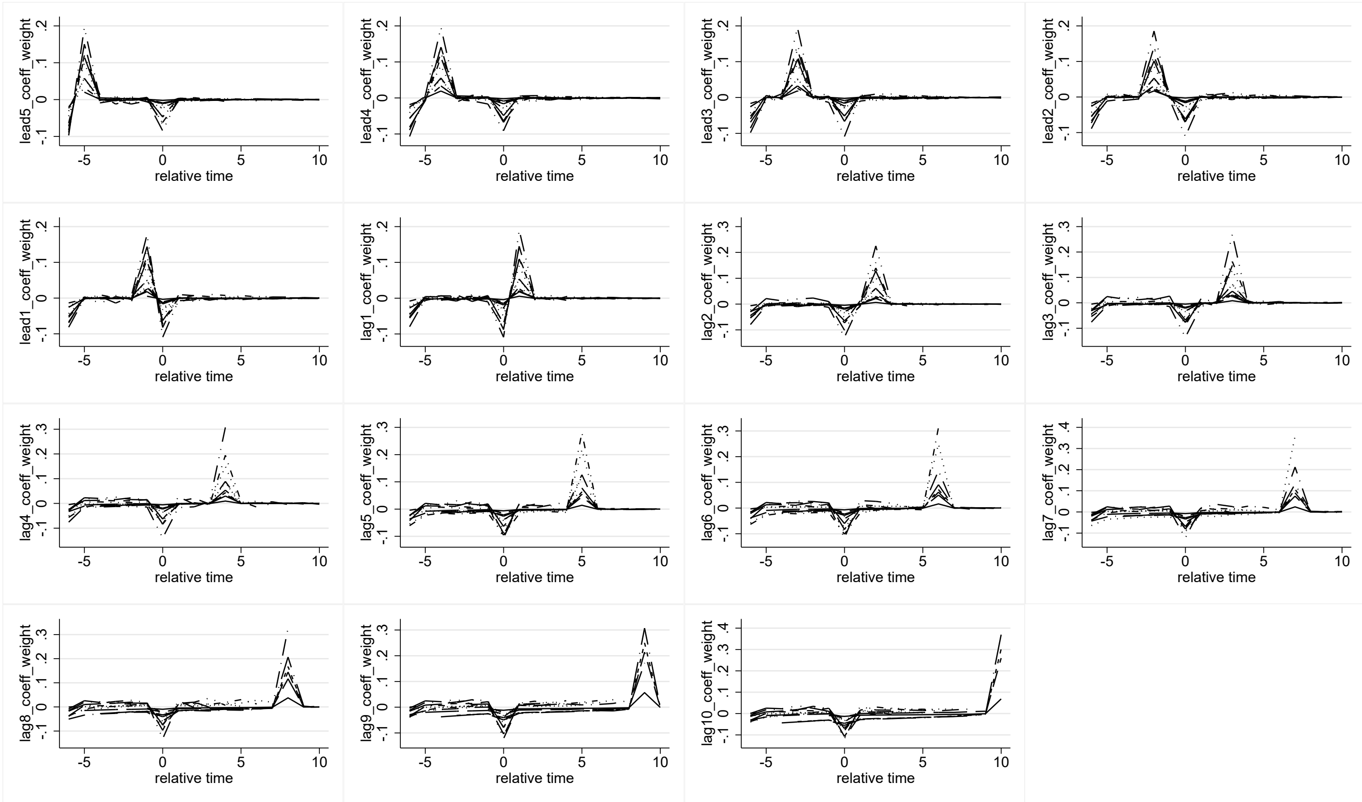
Notes: Data source: IPUMS US Census, 1920, 1930 and 1940.

Figure A.5 — Event-study weights using Sun and Abraham's (2021) interaction weighted estimator



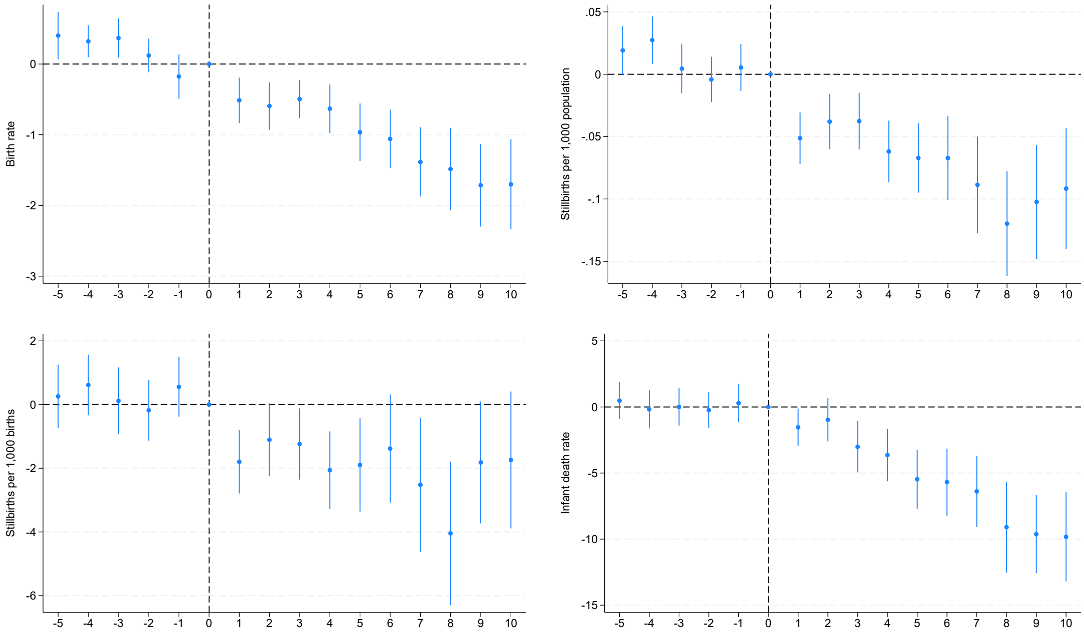
Notes: Data source: U.S. Vital Statistics. The figure shows which cohorts contribute which weight for all lead and lag coefficients estimated in the event-study using Sun and Abraham's (2021) interaction weighted estimator. The sample is restricted to a matched balanced panel of counties for the years 1925 to 1939.

Figure A.6 — Lead and lag specific weights in a dynamic TWFE specification



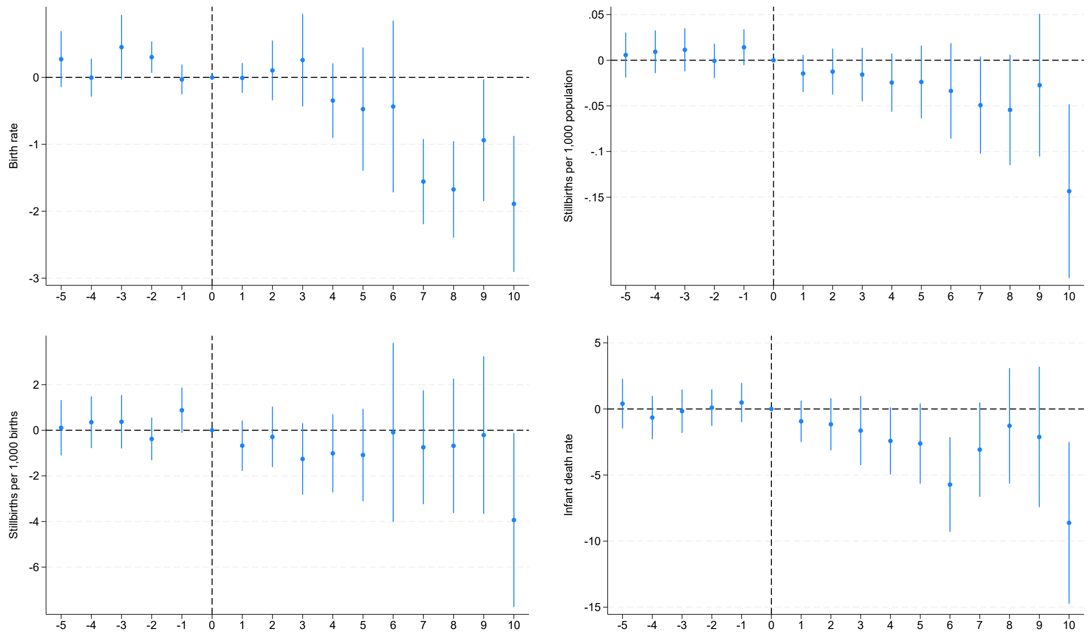
Notes: Data source: U.S. Vital Statistics. The figure shows that in a dynamic two-way fixed effects specification, the estimated event-study coefficients are combinations of differences in trends from their own relative period, from relative periods belonging to other bins included in the specification, and from relative periods excluded from the specification. We employ Sun and Abraham's publicly available Stata package 'eventstudyweights' to estimate the weights. The sample is restricted to a matched balanced panel of counties for the years 1925 to 1939.

Figure A.7 — The impact of birth control clinics in a 20 years panel (Sun and Abraham, 2021)



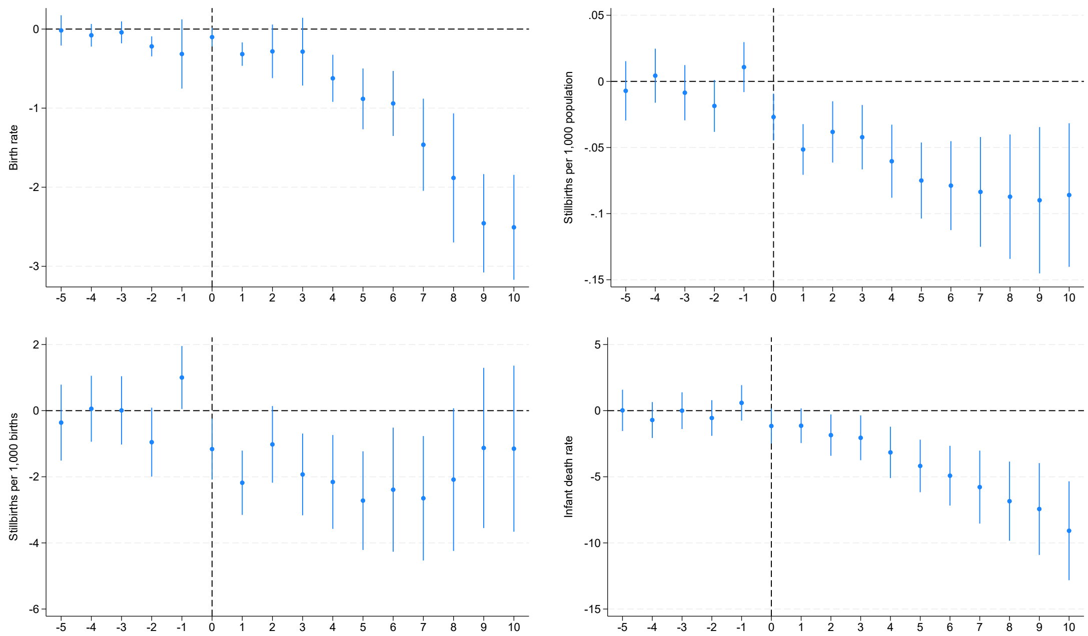
Notes: Data source: U.S. Vital Statistics. The figure shows the dynamic impact of the establishment of the first birth control clinic in a county (or an adjacent county) on the birth rate (births per 1,000 population, upper left panel), stillbirths per 1,000 population (upper right panel), stillbirths per 1,000 births incl. stillbirths (lower left panel) and the infant death rate (infant deaths per 1,000 births, lower right panel). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). The control group consists of never-treated counties. Standard errors are clustered at the county level. The whiskers mark the 95 percent confidence band. The sample is restricted to a matched balanced panel of counties for the years 1920 to 1939.

Figure A.8 — The impact of birth control clinics using the last treated cohort as control group (Sun and Abraham, 2021)



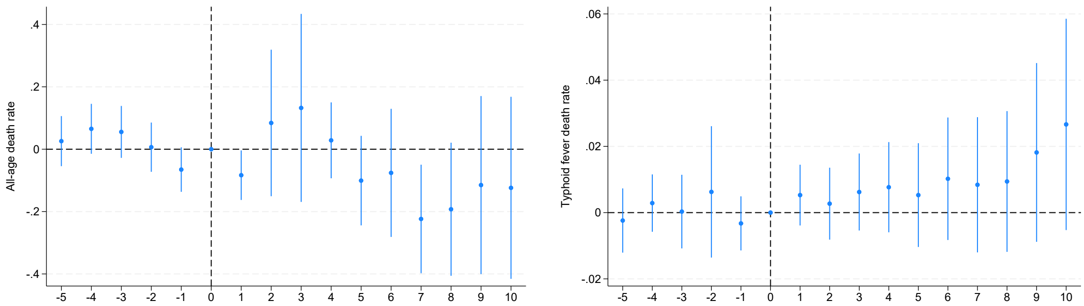
Notes: Data source: U.S. Vital Statistics. The figure shows the dynamic impact of the establishment of the first birth control clinic in a county (or an adjacent county) on the birth rate (births per 1,000 population, upper left panel), stillbirths per 1,000 population (upper right panel), stillbirths per 1,000 births incl. stillbirths (lower left panel) and the infant death rate (infant deaths per 1,000 births, lower right panel). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). The control group consists of the counties that are treated in 1938 and 1939. Standard errors are clustered at the county level. The whiskers mark the 95 percent confidence band. The sample is restricted to a balanced panel of counties for the years 1925 to 1937.

Figure A.9 — The impact of birth control clinics (Callaway and Sant'Anna, 2021)



Notes: Data source: U.S. Vital Statistics. The figure shows the dynamic impact of the establishment of the first birth control clinic in a county (or an adjacent county) on the birth rate (births per 1,000 population, upper left panel), stillbirths per 1,000 population (upper right panel), stillbirths per 1,000 births incl. stillbirths (lower left panel) and the infant death rate (infant deaths per 1,000 births, lower right panel). Treatment effects are derived using the estimator by Callaway and Sant'Anna (2021). The control group consists of never-treated counties. The whiskers mark the 95 percent confidence band. The sample is restricted to a matched balanced panel of counties for the years 1925 to 1939.

FIGURE A.10 — The impact of birth control clinics on all-age total mortality and typhoid fever mortality (Sun and Abraham, 2021)



Notes: Data source: U.S. Vital Statistics and city level causes of death data. The figure shows the dynamic impact of the establishment of the first birth control clinic in a county (or an adjacent county) on the all-age death rate (deaths without infant deaths per 1,000 population) (left panel) and the typhoid fever death rate (typhoid fever deaths per 1,000 population) (right panel). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). The control group consists of never-treated counties or cities, respectively. Standard errors are clustered at the county level. The whiskers mark the 95 percent confidence band. The sample in the left panel is restricted to a matched balanced panel of counties for the years 1925 to 1939. The sample in the right panel is restricted to a balanced panel of cities for the years 1920 to 1937.

Table A.1 — County-level correlates of the roll-out of birth control clinics

	Cty. has BCC by 1940 Charact. of 1920 (I)	Cty. has BCC by 1940 Charact. of 1920 (II)	Cty. has BCC by 1940 Charact. of 1920 (III)	County expos. time Charact. of 1920 (IV)	County expos. time Charact. of 1920 (V)	County expos. time Charact. of 1920 (VI)
Fertility in 1920	-0.036	-0.032	-0.031	-0.233	-0.282	-0.266
	(0.033)	(0.034)	(0.034)	(0.244)	(0.257)	(0.254)
Urban pop. share	0.194	0.163	0.296**	0.898	0.710	1.624*
	(0.122)	(0.123)	(0.129)	(0.914)	(0.919)	(0.952)
Fertility in 1920 × Urban pop. Share	-0.101	-0.105	-0.218	0.849	0.586	-0.830
	(0.151)	(0.152)	(0.160)	(1.133)	(1.134)	(1.176)
Population share living in farm households	-0.153***	-0.135**	-0.101	-1.024***	-0.640	-0.244
	(0.052)	(0.053)	(0.069)	(0.391)	(0.398)	(0.505)
Population living in big city (>100,000)	0.246***	0.225***	0.196*	6.572***	6.200***	2.578***
	(0.076)	(0.077)	(0.102)	(0.573)	(0.575)	(0.750)
Population share women 15-19 years old	-0.411	-0.244	-0.325	-3.058	-0.624	-1.288
	(0.404)	(0.490)	(0.511)	(3.031)	(3.660)	(3.764)
Population share women 20-24 years old	0.059	0.014	-0.134	0.244	1.279	-0.405
	(0.426)	(0.493)	(0.503)	(3.198)	(3.682)	(3.709)
Population share women 25-29 years old	0.404	0.451	0.207	2.461	4.046	1.519
	(0.482)	(0.538)	(0.542)	(3.623)	(4.018)	(3.993)
Population share women 30-34 years old	0.016	0.189	0.027	1.480	4.124	3.135
	(0.542)	(0.621)	(0.627)	(4.069)	(4.642)	(4.622)
Population share women 35-39 years old	0.664	0.695	0.564	4.455	5.401	4.826
	(0.516)	(0.577)	(0.580)	(3.875)	(4.311)	(4.270)
Population share women 40-44 years old	0.987*	1.018	0.829	8.942**	9.684**	9.997**
	(0.588)	(0.650)	(0.655)	(4.413)	(4.854)	(4.822)
Population share women 45-49 years old	1.202**	1.299*	1.112*	7.120	8.952*	9.112*
	(0.605)	(0.663)	(0.675)	(4.546)	(4.957)	(4.976)
Female to male ratio (15-49 years)		-0.045	0.005		-0.540	-0.313
		(0.054)	(0.059)		(0.405)	(0.435)
Share of women 15-49 foreignborn		0.137	0.114		3.350***	2.245***
		(0.108)	(0.113)		(0.809)	(0.831)
Share of women 15-49 literate		0.009	-0.011		0.528	0.665
		(0.137)	(0.143)		(1.021)	(1.052)
Share of women 15-49 black		-0.061	-0.090		-1.269**	-1.469**
		(0.079)	(0.083)		(0.591)	(0.608)
Share of women 15-49 non-white & non-black		-0.086	-0.079		-0.465	-0.570
		(0.147)	(0.148)		(1.095)	(1.093)
Share of women 15-49 in labor force		0.237***	0.241**		2.154***	2.296***
		(0.089)	(0.094)		(0.664)	(0.691)
Share of adults ever married			-0.104			-0.142
			(0.140)			(1.031)
Population size (in 1,000)			0.000			0.005***
			(0.000)			(0.001)
Population share protestant			Ref.			Ref.
Population share catholic			-0.018			0.252
			(0.057)			(0.420)
Share other religion			-0.092			-0.783
			(0.078)			(0.574)
WW1 veterans per male adults in 1930			0.048			-0.359
			(0.161)			(1.189)
Occupational income score			0.005			0.033
			(0.004)			(0.030)
Share manuf. empl. in 1930			-0.074			2.600**
			(0.148)			(1.090)
State-fixed effects	yes	yes	yes	yes	yes	yes
Adj. R2, pseudo R2	3,006	3,006	2,956	3,006	3,006	2,956
Observations	0.276	0.277	0.284	0.074	0.076	0.076

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. In cols. (I)-(III), we regress a binary variable 'having a birth control clinic in the own or adjacent county by 1940' on characteristics observed in 1920. In cols. (IV)-(VI), we regress the time a clinic existed in the own or adjacent county by 1940, again on characteristics observed in 1920 and 1940 respectively. Since for all never-treated counties this time is zero, we use a tobit model. 'Fertility' is measured by the number of a woman's own children below the age of five living in the household.

Table A.2 — The impact of exposure to a birth control clinic on fertility: age patterns

	<20 years	20-24 years	25-29 years	30-34 years	35-39 years	All own childr. in hh	No child at 30	No child at 39
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Years of exp. to BCC (coef. × 100)	-0.695***	-0.851***	-0.339***	-0.132*	-0.112*	-1.909***	0.023	-0.022
	(0.116)	(0.099)	(0.088)	(0.069)	(0.060)	(0.230)	(0.042)	(0.047)
Age FE	yes	yes	yes	yes	yes	yes	yes	yes
Urbanity control	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes
County FE	yes	yes	yes	yes	yes	yes	yes	yes
Year FE × Urbanity	yes	yes	yes	yes	yes	yes	yes	yes
Socio-economic controls	yes	yes	yes	yes	yes	yes	yes	yes
Sample mean of dep. var.	0.444	0.779	0.780	0.603	0.420	1.701	0.257	0.214
R2	0.044	0.051	0.068	0.070	0.076	0.196	0.055	0.058
Observations	2,091,679	8,550,321	11,526,383	11,538,534	11,413,820	45,120,737	11,829,772	6,679,617

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. The table shows OLS regressions. In cols. (I) to (V) the dependent variable is the number of a woman's own children below the age of five living in the household. Col. (VI) uses as a dependant variable the woman's number of children regardless of age living in the household. Cols. (VII) and (VIII) consider women aged 28-32 and 37-39 respectively. For readability, the coefficient associated with the years of exposure to a birth control clinic is muliplied by 100. Socio-economic controls are the literacy status, race, an indicator for being foreign born, an indicator for living in a big city, an indicator for living in a farm household, and the county's religuous composition. Standard errors are clustered at county level.

TABLE A.3 — The impact of exposure to a birth control clinic on fertility, estimates using sub-sammples

	Benchmark (I)	w/t 1940 (II)	Panel (III)	Matched panel (IV)
Years of expos. to BCC (coef. x 100)	-0.403***	-0.784***	-0.429***	-0.421***
	(0.048)	(0.089)	(0.085)	(0.091)
Age FE	yes	yes	yes	yes
Urbanity control	yes	yes	yes	yes
Year FE	yes	yes	yes	yes
County FE	yes	yes	yes	yes
Year FE × Urbanity	yes	yes	yes	yes
Socio-economic controls	yes	yes	yes	yes
Sample mean of dep. var.	0.628	0.681	0.639	0.629
R2	0.093	0.094	0.091	0.091
Observations	45,120,737	28,337,456	28,244,756	25,023,528

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. The table shows OLS regressions. The dependent variable is the number of a woman's own children below the age of five living in the household. For readability, the coefficient associated with the years of exposure to a birth control clinic is muliplied by 100. Socio-economic controls are the literacy status, race, an indicator for being foreign born, an indicator for living in a big city, an indicator for living in a farm household, and the county's religuous composition. Standard errors are clustered at county level. Col. (I) shows the benchmark estimation using the full sample and the specification of Table 2, col. (IV). Col. (II) uses only the Census data of 1920 and 1930. Col. (III) uses only those counties that are part of the panel of counties used in Section 5. Col. (IV) uses only those counties that are part of the matched panel of counties used in Section 5.

Table A.4 — The impact of exposure to a birth control clinic on fertility: alternative specifications

	(I)	(II)	(III)
ln (Years of exposure to BCC +1)	-0.019***		
	(0.003)		
Sine transf. (Years of exposure to BCC)		-0.015***	
		(0.002)	
Years of exposure since county level av. age at marriage			-0.437***
			(0.069)
Age FE	yes	yes	yes
Urbanity control	yes	yes	yes
Year FE	yes	yes	yes
County FE	yes	yes	yes
Year FE × Urbanity	yes	yes	yes
Socio-economic controls	yes	yes	yes
R-squared	0.093	0.093	0.093
Observations	45,120,737	45,120,737	45,104,220

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. The table shows OLS regressions. The dependent variable is the number of a woman's own children below the age of five living in the household. In col. (III) exposure to a birth control clinic is not measured since the age of 15 or the opening of the clinic (whichever came later), but from the counties average age at marriage or the opening of the clinic (whichever came later). For readability, the coefficient associated with the years of exposure to a birth control clinic is muliplied by 100. Socio-economic controls are the literacy status, race, an indicator for being foreign born, an indicator for living in a big city, an indicator for living in a farm household, and the county's religuous composition. Standard errors are clustered at county level.

TABLE A.5 — The impact of exposure to a birth control clinic on fertility by buffer zones

	20km zone (I)	50km zone (II)	50-100km ring (III)
Years of expos. to BCC (coef. × 100)	-0.373***	-0.459***	-0.286***
	(0.053)	(0.045)	(0.064)
Age FE	yes	yes	yes
Urbanity control	yes	yes	yes
Year FE	yes	yes	yes
County FE	yes	yes	yes
Year FE × Urbanity	yes	yes	yes
Socio-economic controls	yes	yes	yes
R2	0.095	0.092	0.092
Observations	31,033,840	43,450,852	25,526,847

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. The table shows OLS regressions. The dependent variable is the number of a woman's own children below the age of five living in the household. For readability, the coefficient associated with the years of exposure to a birth control clinic is muliplied by 100. Socio-economic controls are the literacy status, race, an indicator for being foreign born, an indicator for living in a big city, an indicator for living in a farm household, and the county's religuous composition. In cols. (I) and (II) we drop observations that have access to a birth control clinic within the 20 to 50km and 50 to 100km ring respectively. In col. (III) we drop observations that have access to a birth control clinic within the 50km buffer zone. Standard errors are clustered at the birth cohort by county level.

Table A.6 — The impact of exposure to a birth control clinic on fertility, heterogeneity (coefficients × 100)

	Yes (I)	No (II)	(I) vs. (II) p-value (III)
Black (individual level)	-0.582***	-0.278***	
	(0.056)	(0.052)	(<0.001)
	[0.523]	[0.642]	
Foreign born (individual level)	-0.156	-0.128***	
	(0.186)	(0.047)	(0.882)
	[0.719]	[0.616]	
Big cities (cities >100,000 pop.)	-0.396***	-0.608***	
	(0.088)	(0.055)	(0.041)
	[0.504]	[0.690]	
Catholic share at county level above median share	-0.384***	-0.368***	
	(0.075)	(0.049)	(0.864)
	[0.590]	[0.667]	
County vote in 1920 majority Democrats	-0.561***	-0.323***	
	(0.102)	(0.053)	(0.038)
	[0.652]	[0.620]	
County vote in 1932 majority Democrats	-0.457***	-0.162**	
	(0.070)	(0.074)	(0.004)
	[0.635]	[0.608]	
Southern states	-0.618***	-0.327***	
	(0.123)	(0.051)	(0.029)
	[0.662]	[0.617]	
Above median annual retail sales growth (1929-39)	-0.493***	-0.426***	
	(0.075)	(0.074)	(0.521)
	[0.648]	[0.628]	
Above median annual growth of social spending (1933-39)	-0.378***	-0.544***	
	(0.099)	(0.060)	(0.153)
	[0.628]	[0.652]	
Above median unemployment in 1930	-0.310***	-0.409***	
	(0.060)	(0.064)	(0.258)
	[0.574]	[0.679]	
Above median WWI veterans in 1930	-0.374***	-0.519***	
	(0.071)	(0.089)	(0.201)
	[0.566]	[0.689]	

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940, NHGIS county level data, and election data taken from Robinson (1934). Counties heavily affected by the Great Depression are counties with a below median retail sales growth from 1929 to 1939 (see Fishback et al., 2005). Counties with a large prevalence of welfare benefit payments in the 1930s are defined as having above median social spending growth over the period 1933-1939 (see Fishback et al., 2005). The sample, specification and controls used are the same than those used in Table 2, col. (IV). The dependent variable is the number of a woman's own children below the age of five living in the household. Each coefficient comes from a different regression and, for readability, is multiplied by 100. Group-specific sample means in brackets.

Table A.7 — Matching treatment and control counties

	Probit		Sample before matching				Matched sample			
	Treated		Treated		Control		Treated		Control	
	coeff	std.err.	mean	std.dev.	mean	std.dev.	mean	std.dev.	mean	std.dev.
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)
Birth rate (1925)	0.036***	(0.008)	22.966	(11.396)	21.548	(7.742)	22.708	(11.045)	21.719	(4.926)
Stillbirths per 1,000 births (1925)	0.005**	(0.002)	35.904	(14.861)	31.91	(16.825)	36.125	(14.761)	34.386	(17.533)
Infant death rate (1925)	0.002	(0.002)	70.445	(21.875)	66.519	(40.503)	70.226	(21.901)	66.805	(24.276)
All-age mortality rate (1925)	0.074***	(0.014)	10.068	(3.272)	8.596	(4.544)	9.939	(2.898)	9.387	(3.100)
Urban population share	-0.704***	(0.231)	0.278	(0.278)	0.159	(0.215)	0.276	(0.273)	0.195	(0.230)
Population living in farm households	-0.886***	(0.294)	0.433	(0.248)	0.563	(0.216)	0.434	(0.246)	0.510	(0.220)
Population living in big city (>100,000)	-0.483	(0.586)	0.036	(0.160)	0.001	(0.032)	0.036	(0.161)	0.003	(0.049)
Population share women 15-19 years old	-5.119**	(2.252)	0.045	(0.017)	0.046	(0.023)	0.045	(0.016)	0.046	(0.021)
Population share women 20-24 years old	-2.978	(2.145)	0.041	(0.017)	0.040	(0.020)	0.041	(0.016)	0.041	(0.018)
Population share women 25-29 years old	-3.926*	(2.370)	0.038	(0.015)	0.038	(0.019)	0.038	(0.015)	0.037	(0.016)
Population share women 30-34 years old	-3.104	(2.558)	0.034	(0.013)	0.034	(0.019)	0.034	(0.013)	0.034	(0.017)
Population share women 35-39 years old	0.974	(2.811)	0.033	(0.012)	0.030	(0.016)	0.033	(0.012)	0.031	(0.015)
Population share women 40-44 years old	1.803	(2.684)	0.028	(0.013)	0.026	(0.016)	0.028	(0.013)	0.028	(0.016)
Population share women 45-49 years old	4.310	(2.782)	0.025	(0.012)	0.023	(0.015)	0.025	(0.012)	0.025	(0.016)
Female to male ratio (15-49 years)	0.623**	(0.245)	0.973	(0.202)	0.950	(0.218)	0.973	(0.198)	0.963	(0.219)
Share of women 15-49 foreignborn	0.653	(0.431)	0.088	(0.105)	0.072	(0.097)	0.089	(0.106)	0.069	(0.087)
Share of women 15-49 literate	-2.032***	(0.760)	0.954	(0.057)	0.967	(0.063)	0.955	(0.057)	0.958	(0.070)
Share of women 15-49 black	-1.332***	(0.368)	0.073	(0.135)	0.062	(0.155)	0.070	(0.133)	0.072	(0.165)
Share of women 15-49 non-white & non-black	0.459	(0.802)	0.022	(0.045)	0.018	(0.046)	0.021	(0.045)	0.023	(0.053)
Share of women 15-49 in labor force	0.045	(0.408)	0.195	(0.103)	0.157	(0.108)	0.194	(0.102)	0.172	(0.108)
Share of adults ever married	-1.939***	(0.570)	0.753	(0.064)	0.763	(0.078)	0.753	(0.064)	0.754	(0.075)
Population size (in 1,000)	0.007***	(0.001)	55.675	(120.64)	19.106	(28.369)	56.421	(121.73)	23.283	(40.291)
Population share catholic	-0.908***	(0.220)	0.231	(0.219)	0.217	(0.209)	0.235	(0.219)	0.214	(0.202)
Share other religion	-0.739***	(0.231)	0.137	(0.130)	0.165	(0.180)	0.137	(0.131)	0.152	(0.158)
WW1 veterans per male adults in 1930	-1.261*	(0.716)	0.078	(0.046)	0.080	(0.055)	0.078	(0.041)	0.080	(0.057)
Occupational income score	0.013	(0.018)	19.888	(3.591)	18.302	(3.519)	19.89	(3.570)	18.923	(3.450)
Share manuf. empl. in 1930	1.837***	(0.565)	0.084	(0.087)	0.043	(0.062)	0.084	(0.087)	0.056	(0.070)
Observations in 1925	1,773		905		916		888		392	
Pseudo R2	0.154									

Notes: Data sources: U.S. Vital Statistics. The table shows the results of a probit model where the outcome variable is a dummy variable indicating whether a birth control clinic was established in a county or a neigboring county prior to 1940 (col. I and II) as well as a comparison of treated and control counties in the balanced panel 1925-1939 (col. III to VI) and a comparison of treated and control counties in the matched balanced panel 1925-1939 (col. VII to X). All predictor variables are measured in 1920 unless indicated otherwise.

Table A.8 — Vital statistics: Descriptives

	Balanced panel 1925-1939, matched (I)	Balanced panel 1920-1939, matched (II)	Balanced panel 1925-1937, no never-treated (III)	Unbalanced panel 1920-1939 (IV)	Balanced panel 1920-1937 (V)
Birth rate	19.392	20.685	19.752	20.338	
	(7.694)	(10.164)	(9.416)	(8.310)	
Stillbirth rate (per 1,000 pop)	0.672	0.684	0.702	0.705	
	(0.403)	(0.386)	(0.404)	(0.444)	
Stillbirth rate (per 1,000 births)	32.822	32.282	33.805	33.325	
	(15.555)	(14.136)	(14.895)	(18.382)	
Infant death rate	59.238	62.378	61.868	62.054	
	(21.270)	(21.653)	(20.854)	(27.173)	
All-age death rate	9.954	10.046	10.270	9.244	
	(3.093)	(2.952)	(3.526)	(3.682)	
Puerperal death rate					0.587
					(0.368)
Typhoid fever death rate					0.066
					(0.104)
Distinct counties	1,280	924	907	3,011	
Distinct cities					315
Obs.	19,200	18,480	11,791	48,515	5,670

Notes: Data sources: U.S. Vital Statistics. The table shows the means and standard deviations (in parentheses) of fertility and mortality variables in the matched balanced county panel 1925-1939 (main sample, col. I), the matched balanced county panel 1920-1939 (col. II), the balanced county panel 1925-1938 without never-treated counties (col. III), the unbalanced county panel 1920-1939 (col. IV) as well as the balanced city panel 1920-1937 (col. V). The birth rate is defined as births per 1,000 population, the stillbirth rate is defined as stillbirths per 1,000 population, or 1,000 births incl. stillbirths, respectively, the infant death rate is defined as infant deaths per 1,000 births, the all-age death rate is defined as deaths without infant deaths per 1,000 population. The puerperal death rate is defined as puerperal deaths per 1,000 women aged 15-49, the typhoid fever death rate is defined as typhoid fever deaths per 1,000 population.

Table A.9 — Validity and subsample checks

	Birth rate (I)	Stillbirths per 1,000 population (II)	Stillbirths per 1,000 births (III)	Infant death rate (IV)
Urbanization				
Urban counties	-0.880***	-0.036***	-1.270*	-1.305
	(0.183)	(0.014)	(0.765)	(0.842)
	3,500	*3,500*	*3,500*	*3,500*
Non-urban counties	-0.406**	-0.030***	-0.973**	-1.843**
	(0.170)	(0.010)	(0.497)	(0.792)
	12,246	*12,247*	*12,246*	*12,187*
Welfare benefits				
Social spending below median	-0.565***	-0.037***	-1.128**	-1.160
	(0.123)	(0.012)	(0.564)	(0.739)
	7,878	*7,878*	*7,878*	*7,865*
Social spending above median	-0.656***	-0.031***	-0.937	-2.555***
	(0.223)	(0.012)	(0.585)	(0.958)
	7,868	*7,869*	*7,868*	*7,822*
Social spending growth below median	-0.572***	-0.033**	-0.684	-0.905
	(0.214)	(0.013)	(0.616)	(0.913)
	7,886	*7,887*	*7,886*	*7,848*
Social spending growth above median	-0.614***	-0.036***	-1.300**	-2.805***
	(0.151)	(0.010)	(0.523)	(0.811)
	7,859	*7,859*	*7,859*	*7,838*
Great Depression				
Retail sales growth above median	-0.775***	-0.035***	-1.203**	-2.322**
	(0.159)	(0.012)	(0.591)	(0.923)
	7,880	*7,881*	*7,880*	*7,853*
Retail sales growth below median	-0.441**	-0.030***	-0.733	-1.781**
	(0.202)	(0.011)	(0.537)	(0.803)
	7,866	*7,866*	*7,866*	*7,834*
Unemployment 1930 below median	-0.776***	-0.035***	-0.604	-2.236***
	(0.179)	(0.009)	(0.510)	(0.822)
	7,874	*7,874*	*7,880*	*7,854*
Unemployment 1930 above median	-0.219	-0.033**	-1.744***	-0.573
	(0.173)	(0.014)	(0.648)	(0.871)
	7,863	*7,864*	*7,863*	*7,824*

Notes: Data sources: U.S. Vital Statistics. The table shows the impact of the establishment of the first birth control clinic in a county (or an adjacent county) on the birth rate (births per 1,000 population) (col. I), stillbirths per 1,000 population (col. II), stillbirths per 1,000 births incl. stillbirths (col. III), and the infant death rate (infant deaths per 1,000 births) (col. IV) for various subsamples. Urban counties are counties where more than 50 percent of households live in urban areas by 1940 according to the Census. Counties with a large prevalence of welfare benefit payments in the 1930s are defined as having above median per capita public works and relief spending over the period 1933-1939, or having above median per capita public works and relief spending in the period 1933-1939 as compared to 1933-1935 (see Fishback et al., 2005). Counties heavily affected by the Great Depression are counties with a below median retail sales growth from 1929 to 1939 (see Fishback et al., 2005) or counties with an above median unemployment rate in 1930. Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). We bin all post-treatment years into a single indicator. The control group consists of never-treated counties. Standard errors reported in parentheses are clustered at the county level. The number of observations is printed in italics. The sample is restricted to a matched balanced panel of counties for the years 1925 to 1939.

A.2. Data sources

Birth control clinic data set

Digitized data set of birth control clinics established in the U.S. before 1940 based on Hajo (2010). Hajo gathered the information from various sources, including historical issues of the Birth Control Review and press archives. For each clinic, the dataset provides information on the county in which it was located, the year of its establishment, and, if applicable, the year of its closure. The data set encompasses a total of 639 birth control clinics, which are geographically dispersed across 44 states.

Press coverage (Newspapers.com)

Newspapers.com is a online newspaper archive. We recorded the number of newspaper articles mentioning the term "birth control" between 1910 and 1940.

US Census Data (IPUMS USA)

Census 1920 Women's age, race, nativity, literacy, farm status, urbanity and size of place, number of own children below 5 in household, number of all own children in household, and labor force participation (all individual level). Population size, populage age and sex composition, sex ratio, urban population share, socio-demographic population composition, and occupation income score (all aggregated at county level).

Census 1930 Women's age, race, nativity, literacy, farm status, urbanity and size of place, number of own children below 5 in household, number of all own children in household, and labor force participation (all individual level). Age at marriage and World War I veteran status (all aggregated at county level).

Census 1940 Women's age, race, nativity, education, farm status, urbanity and size of place, number of own children below 5 in household, number of all own children in household, and labor force participation (all individual level).

For more details see: Ruggles, S., Flood, S., Goeken, R., Schouweiler, M. and Sobek, M. (2022). IPUMS USA: Version 12.0. Minneapolis, MS: IPUMS.

National Historical GIS (NHGIS) (IPUMS USA)

Census of Religious Bodies: Religious Bodies Data Set 1906-1936 (county level) Percentage of different religious groups.

1930 Census: Population, Agriculture & Economic Data (county level) Percentage of unemployed workers, percentage of manufacturing workers.

For more details see: Manson, S., Schroeder, J., Van Riper, D., Kugler, T., and Ruggles, S. IPUMS National Historical Geographic Information System: Version 17.0 [dataset]. Minneapolis, MN: IPUMS. 2022.

County Longitudinal Template, 1840-1990 (Horan-Hargis)

The County Longitudinal Template is a tool that allows to account for temporal changes in the geographic boundaries of counties in the United States due to the split of counties, the merge of counties or changes in the boundaries of counties. These data provide a decade-by-decade account of the administrative status of each county, starting in 1990 and tracing each census period back through 1840. We use this template to work with constant county boundaries over the period 1920-1940. For more details, see Horan, Patrick M., and Peggy G. Hargis. County Longitudinal Template, 1840-1990 [Computer file]. ICPSR version. Athens, GA: Patrick M. Horan, University of Georgia, Dept. of Sociology/Statesboro, GA: Peggy G. Hargis, Georgia Southern University, Dept. of Sociology and Anthropology [producers], 1995. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1995.

List of counties and adjacent counties

County adjacency 2010 set up by Jean Roth (jroth@nber.org), 8 May 2017. NBER URL: $http : //www.nber.org/data/county_adjacency.html$.
Source Page: $https : //www.census.gov/geo/reference/county - adjacency.html$.
Source File URL: $http : //www2.census.gov/geo/docs/reference/county_adjacency.txt$.
Source Text File: $http : //www2.census.gov/geo/docs/reference/county_adjacency.txt$.

Fishback *et al.* (2005) data

County level data on retail sales (1929 and 1939) and per capita public works and relief spending (1933-1935 and 1933-1939).

For more details see: Fishback, P. V., Horrace, W. C., & Kantor, S. (2005). Did New Deal Grant Programs Stimulate Local Economies? A Study of Federal Grants and Retail Sales during the Great Depression. *The Journal of Economic History*, 65(1), 36–71.

Election data

County level voting outcomes of presidential elections of the years 1920 and 1932.

For more details see: Robinson, E. (1934). The Presidential Vote, 1896–1932. Stanford: Stanford University Press.

Vital statistics

U.S. County-Level Natality and Mortality Data, 1915-2007 Live births (exclusive of stillbirths), infant deaths (i.e., deaths of children below the age of one exclusive of stillbirths), total number of deaths by all ages, and total population (all 1920-1939).

For more details see: Bailey, M.J., Clay, K., Fishback, P., Haines, M., Kantor, S., Severnini, E. and Wentz, A. (2016). U.S. County-Level Natality and Mortality Data, 1915–2007. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

NHGIS (IPUMS) County level data on stillbirths (1922 to 1939).

For more details see: Manson, S., Schroeder, J., Van Riper, D., Kugler, T. and Ruggles, S. (2022). IPUMS National Historical Geographic Information System: Version 17.0. Minneapolis, MS: IPUMS.

Causes of death data

Vital statistics of the United States annual volumes 1920-1937 (city level) Puerperal deaths, typhoid fever deaths. Digitized by Ager *et al.* (2024)

For more details see: Ager, P., Feigenbaum, J. J., Hansen, C. W. and Tany, H. R. (2024). How the other half died: Immigration and mortality in US cities. *The Review of Economic Studies*, 91 (1), 1–44.