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*

November 7, 2022

Abstract

Margaret Sanger established the first birth control clinic in New York in 1916. From the mid-1920s, "Sanger clinics" spread over the entire U.S. Combining newly digitized data on the roll-out of clinics, full-count Census data, and administrative vitality statistics, we find that being exposed to a clinic throughout the fertile ages lowers women's birth rates by 12–15%. Moreover, birth control clinics reduce stillbirths and infant mortality as well as all-age mortality, driven by less puerperal deaths. Thus, birth control clinics particularly reduce births characterized by significant health risks. Further suggestive evidence points towards positive effects on female employment.

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

[†] *Corresponding author:* Michael Grimm, School of Business Administration, Economics, and Information Systems, University of Passau, Innstrasse 27, 94032 Passau, Germany, email: michael.grimm@uni-passau.de. For helpful discussions and comments we would like to thank participants at the GDE Conference 2022 in Stuttgart-Hohenheim, the GlaD-DENeB Seminar in Berlin as well as the Berlin Applied Micro Seminar, and seminars participants at the University of Connecticut, 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, 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 codified the Comstock Act, which prohibited the mailing, shipping, or importation of articles, drugs, medicines, and printed materials of obscenities, which also included any article or information designed or intended for the prevention of conception or procuring of abortions (Engelman, 2011, p. 15ff). It took until 1937 that birth control was embraced by the public and the medical profession, and finally got legalized. Shortly thereafter, the Planned Parenthood Federation of America was established and a nationwide network of family planning centers was rolled out (Engelman, 2011, p. 179ff). Although the legalization of birth control in 1937 marked an important victory of the birth control movement over a conservative morality movement, advertising and sales bans of contraceptives remained in place in many U.S. states until the end of the 1960s (Bailey, 2010).

Margaret Sanger was one of the leading figures of the birth control movement. On 16 October 1916, she opened the United States' first birth control clinic in New York, Brownsville, Brooklyn—despite strong opposition (Engelman, 2011, p. 77). In this clinic, women received advice on effective contraception, any sexual matters and hygiene, with the ultimate goal of improving women's and children's health and to relief poverty among working class women. Although this clinic was shut down by the police ten days after it had opened, 'Sanger clinics' started to spread over the entire U.S. from 1923 onward (Engelman, 2011, p. 153). By 1940, there were more than 600 of them (Hajo, 2010, p. 21). Despite the historical relevance of Margaret Sanger's birth control movement, the causal impact of Sanger's birth control clinics has so far not been assessed. Filling this gap should enhance our understanding of the historical fertility transition and its socio-economic effects in the U.S.

In this paper, we assess the causal impact of birth control clinics on fertility, stillbirths, infant mortality, and all-age mortality (without infant deaths). To this end, we digitized detailed data on the roll-out of birth control clinics across U.S. counties from 1916 to

1940. We combine this unique data with full-count Census data of 1920, 1930 and 1940, yearly administrative vital statistics at the county level, and administrative causes of death data at the city level. The Census data allow us to investigate the effects of birth control clinics on fertility. Using the administrative vital statistics, we re-estimate the effect of birth control clinics on fertility and use stillbirths, infant mortality and all-age mortality as additional outcomes. We expect that Sanger's birth control clinics satisfied a demand for contraception and hence increased birth spacing and lowered fertility. An increase in birth spacing can in turn reduce the incidence of fetal and infant deaths. This is because short interpregnancy intervals might mean inadequate time to recover from the insufficient repletion of maternal folate resources, increased risk of vertical transmission of infections to the fetus, or transmission of infections among siblings (Conde-Agudelo *et al.*, 2012). Lower fertility might also have relaxed households' budget constraints and, through this channel, increased health investments per child. Finally, more time between pregnancies and less births might have caused fewer maternal deaths, especially those due to puerperal septicemia or other puerperal causes and due to fewer unprofessional abortions.

In the first part of the empirical analysis, we use full-count Census data and draw on differences in the length of exposure to birth control clinics conditional on birth cohort by state fixed effects and a set of socio-economic individual and county level controls. We find statistically significant and economically meaningful effects on fertility. Being exposed to a birth control clinic throughout the entire fertile period from age 15 to 39 reduces fertility by 12-15%. Event-study analyses provide evidence for the validity of the key identifying assumption. The effects are confirmed in alternative specifications in which we control for county and year fixed effects, use Pseudo Poisson Maximum Likelihood estimators, or restrict the sample to big city counties. Moreover, we follow Altonji *et al.* (2005) and Oster (2019) to show that our results are robust to conventional assumptions about selection on unobservables. Finally, we find some suggestive evidence for positive knock-on effects on female labor supply.

In the second part of the empirical analysis, we use yearly county-level administrative 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 Sun and Abraham (2021)'s interaction weighted estimator. We confirm the negative effects of birth control clinics on fertility. Moreover, we find that birth control clinics reduce stillbirths as well as infant mortality. Insignificant and small pre-treatment coefficients corroborate the validity of the common trends assumption. Complimentary city-level data provide suggestive evidence that birth control clinics reduce puerperal deaths. These findings suggest that birth control clinics particularly reduce births that are characterized by significant health risks. Overall, we interpret these findings as the consequence of a supply side policy that relaxed constraints on the demand side, i.e., in a Beckerian framework (Becker, 1981) birth control clinics reduced the cost of birth spacing and hence having fewer children.

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 (see, e.g. Bourne Wahl, 1992; Steckel, 1992; Haines, 2000; Greenwood and Seshadri, 2002; Hacker, 2003; Haines and Hacker, 2006; Jones and Tertilt, 2006; Curtis White, 2008; Bleakley and Lange, 2009; Bailey and Collins, 2011; Guinnane, 2011; Wanamaker, 2012; Aaronson *et al.*, 2014; Bailey *et al.*, 2014; Hansen *et al.*, 2018; Bailey and Hershbein, 2018; Beach and Hanlon, forthcoming; Ager *et al.*, 2020; Grimm, 2021). 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) in the U.S. The situation in the early 20th century differs from today's low income country context. At the time we look at, people were poorly informed about contraceptives; even the medical profession was rather uninformed and not trained in these issues. Moreover, contraceptives were hardly accessible through the market and largely illegal; modern highly effective contraceptives did not yet exist.

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 estimation results using the Census data, while Section Section 5 does the same using the administrative data. 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. Within 50 years, the total fertility rate declined from about 5.5 to 3.5 (Jones and Tertilt, 2006; Bailey and Hershbein, 2018). Most couples used withdrawal and abstinence as well as extended breastfeeding to space births (Engelman, 2011, p. 4; Reed, 1978, p. 3; David *et al.*, 2007), but also abortion and—as a method of last resort especially among the very poor—infanticide. The second half of the 19th century also witnessed the emergence of the first rudimentary contraceptive devices such as condoms, vaginal sponges, pessaries, 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, p. 10). At the same time, marriage manuals rapidly diffused that explicitly considered that couples could and should enjoy sexual pleasure without procreation (Knowlton, 1832; Owen, 1876).

The religious community as well as the American Medical Association (AMA), and later a larger morality movement, disapproved this development. Anthony Comstock, a leading political figure of the Christian morality movement, associated contraception with illicit sex and pornography. In 1873, he succeeded in making the U.S. Congress codify the Comstock Act, which outlawed the interstate mailing, shipping, or importation of articles, drugs, medicines, and printed materials of obscenities; this also applied to marriage manuals as well as anything used for the prevention of conception or advertising for it (Engelman, 2011, p. 15). The implementation of the Comstock Act varied across states; yet, it was mostly strictly enforced (Bailey, 2010). People not complying with the law were fined or even arrested, and the respective goods confiscated. Between 1860 and 1890, most states also introduced strict anti-abortion laws, which slowed down the fertility decline (Engelman, 2011, p. 13; Lahey, 2014a; Lahey, 2014b).

2.2. The birth control movement

In opposition to these restrictive interventions, the birth control movement and activism for human sexuality emerged at the beginning of the 20th century.[1] A leading figure of this movement was Margaret Sanger, a nurse whose mother had been through 18 pregnancies in 22 years and died at age 50 of tuberculosis and cervical cancer. She was concerned about the hardship brought about by many births 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 universally available (McCann, 1994, p. 210). On a trip to the Netherlands, Sanger was especially impressed by the low infant and maternal mortality rates in the country that were not least attributed to birth control clinics providing contraceptive advice to women (McCann, 1994, p. 59, p. 211; Hajo, 2010, p. 11).

On 16 October 1916, Margaret Sanger opened the United States’ first birth control clinic (or center for contraceptive instruction) in New York, Brownsville, Brooklyn. In the Brownsville clinic, which was sponsored by wealthy supporters, women received advice on effective contraception, any sexual matters and hygiene. To provide women with pessaries and other contraceptive devices, the clinics had to illegally import these goods. If considered absolutely necessary, women were also referred to doctors for therapeutic abortions, although Sanger was generally against abortions because of the associated

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

health risks. Moreover, supporting abortions would have made the acceptance by the morality movement even more difficult (Engelman, 2011, p. 76; Hajo, 2010, p. 13). Sanger was convinced that pessaries ("occlusive diaphragm"), at a cost of one to two dollars, were the best female contraceptive method available at that time and that personal instruction and examination as well as follow-up visits were essential in order to guarantee their effectiveness. In her view, only birth control clinics could provide such services (McCann, 1994, p. 59; Hajo, 2010, p. 67).[2] The Brownsville clinic was advertised by leaflets in Yiddish, Italian, and English and by many press releases (Reed, 1978, p. 106). Despite more than 480 women attending the clinic on its first days, the clinic was shut down by the police ten days after it had opened (Engelman, 2011, p. 78ff; Hajo, 2010, p. 23). Margaret Sanger was sentenced to 30 days in jail.

In the following years, Sanger established the Birth Control League of New York, whose activism included mailings, conferences, lectures, exhibits and lobbyism for legislative change. Similar leagues followed in other cities. By 1917, there were more than 30 birth control organizations spread over the U.S. In 1921, Sanger set up the American Birth Control League, the predecessor of the Planned Parenthood Federation of America (Engelman, 2011, p. 129). By 1924, the league boasted more than 27,000 members and 10 branches. Moreover, Sanger developed the Birth Control Review, a journal devoted to the birth control movement and mainly distributed through street sales (Engelman, 2011, p. 99).

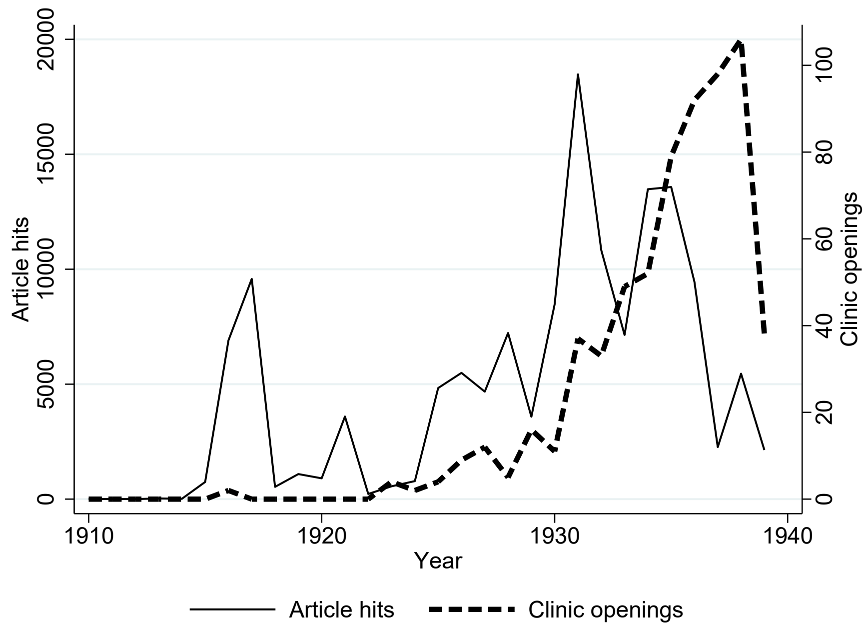
Sanger repeatedly faced criminal accusations and spent short periods of time in jail. She traveled to give speeches, to set up alliances and to mobilize financial resources to fund her activities. Her speeches were regularly interrupted or stopped by the police. Several times, she visited European countries that were already more advanced in her cause. Over time, she slightly adapted her strategy in her fight for the legalization of contraception. Rather than emphasizing women's right for self-determination, she increasingly referred to the health aspects of child spacing—both for mothers and children. With this strategy, she hoped to get doctors and hospitals on board and to avoid conflicts with the law (McCann, 1994, p. 60; Engelman, 2011, p. 101). Margaret Sanger's ultimate objective was to organize birth control clinics nationwide (McCann, 1994, p. 60).

The country-wide roll-out of Sanger's birth control clinics started to take off in 1923 (Engelman, 2011, p. 153) and was facilitated by waning opposition against contraception. Following World War I, many U.S. servicemen were diagnosed with sexually transmitted diseases; as a result, condoms now became an issue of public health (Engelman, 2011, p. 108). There were two types of birth control clinics: independent clinics, such as the Brownsville clinic, typically operated by local birth control organizations, and clinics

[2] McCann (1994, p. 60) 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. A setting in which to evaluate the safety and effectiveness of the various types of pessaries was needed as well because none of them had ever been systematically tested."

embedded in hospitals (Hajo, 2010, p. 19ff). Independent clinics were mostly run by female activists and a male medical doctor (in part time), often the spouse of an activist (Hajo, 2010, p. 73ff). The clinics continuously professionalized and were increasingly staffed with trained practitioners. These were trained either in the clinics themselves or in Sanger's Clinical Research Bureau in New York, which was particularly set up for this purpose. The Clinical Research Bureau was the first legal birth control clinic in the United States, and rapidly grew into the leading contraceptive research center in the world (Reed, 1978, p. 116; Hajo, 2010, p. 13). Between 1916 and 1940 more than 650 "Sanger clinics" were founded in the U.S (Hajo, 2010, p. 21). As Figure 1 shows, 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. A first spike is clearly visible when the Brownsville clinic opened in 1916. Also in the following years, the spikes in the roll-out are correlated with the spikes in press coverage. In the second half of the 1930s, newspaper coverage decreased as opposition vaned and birth control clinics became more and more accepted in the U.S. society.

FIGURE 1 — 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.

Birth control clinics provided their services regardless of a woman's ability to pay.[3] Yet, for exactly this reason, they often incurred large budgetary deficits. Consequently, some clinics were not sustainable and had to close again (McCann, 1994, p. 61). Clinics

[3]The costs for an average patient were around USD 6.50, but USD 5 was typically the highest fee charged (Reed, 1978, p. 116).

targeted and admitted only married women that typically already had several children as the main intention was to allow women to space births but not to prevent births all together. Admission requirements were usually directly checked at the entrance of a clinic before women were passed on to the nurse or doctor in charge (Hajo, 2010, p. 49ff). In the beginning, Afro-American women were not targeted as segregation was still strict at that time. Yet, Afro-Americans increasingly set up their own clinics following the same model. Later, segregation started to vanish and clinics served both White and Afro-American women (Hajo, 2010, p. 83ff). Over the entire period, Afro-Americans made up about 11% of all patients, which roughly corresponded to their population share (Hajo, 2010, p. 113).

Eventually, in 1937, the U.S. Court legalized birth control. It was then embraced by the public and the medical profession. Also the American Medical Association's Committee on Contraception tentatively approved contraception (McCann, 1994, p. 217). Still, in many U.S. states advertising and sales bans in relation to contraception remained in place until the end of the 1960s (Bailey, 2010). In 1939, the American Birth Control League merged with Sanger's clinic to become the Birth Control Federation of America. At that point, Sanger became honorary president and relinquished active leadership. By 1942, the name was changed to Planned Parenthood Federation of America. Margaret Sanger raised funds from Katharine Dexter McCormick and other sources which supported the development of the pill. In 1966, she died in Tucson, Texas. Despite the praise she still receives today for her activism for women's rights, birth control and planned parenthood, she is also criticized for her closeness to the eugenics and her support for forced sterilization (Weisbord, 1973; Engelman, 2011, p. 130ff).

3. Data

3.1. Birth control clinics

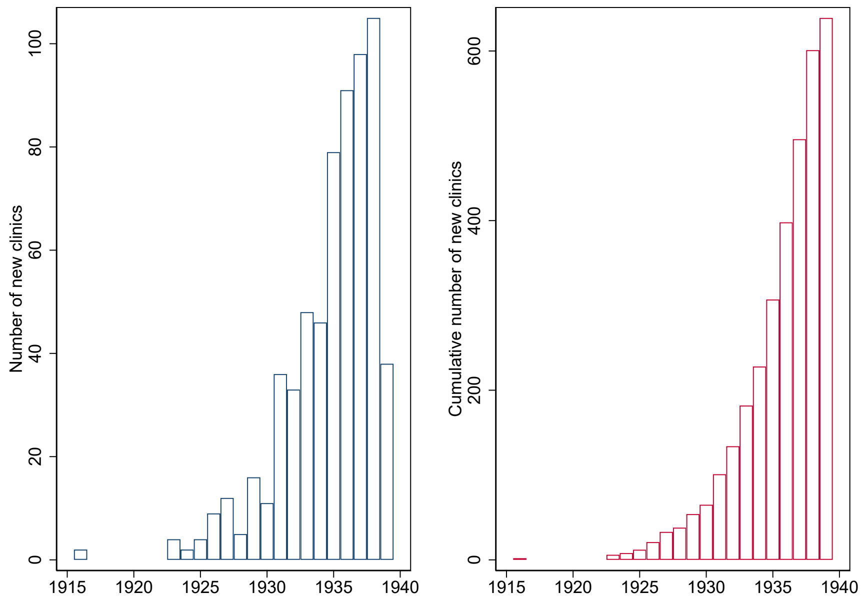
To measure exposure to a birth control clinic, we use a complete inventory of birth control clinics that were established in the U.S. before 1940. We digitized the data collected by Hajo (2010) relying on information published in historical issues of the Birth Control Review, press archives, and many other sources. For each clinic, we have information on the name, the sponsor, the location, the year it was opened as well as, if applicable, the year it was closed. In total, the data set includes 639 birth control clinics distributed across 44 states.[4]

Figure 2 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 new clinics. The first birth control clinic opened in New York, Brownsville, Brooklyn, in 1916; a second one in the same year in

[4]We exclude the states of Alaska and Hawaii; there was no birth control clinic in Alaska.

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 1935, 307 birth control clinics had been established in the U.S. By 1939, the last year of our period of observation, this number had increased to 639. Note that some clinics had to close again after opening. Still, in our empirical analysis, we use the year of the opening as the treatment variable. The reason is that once a clinic had been operating it could have had a lasting effect as the information had been spread and norms might have changed.

FIGURE 2 — Birth control clinics over time



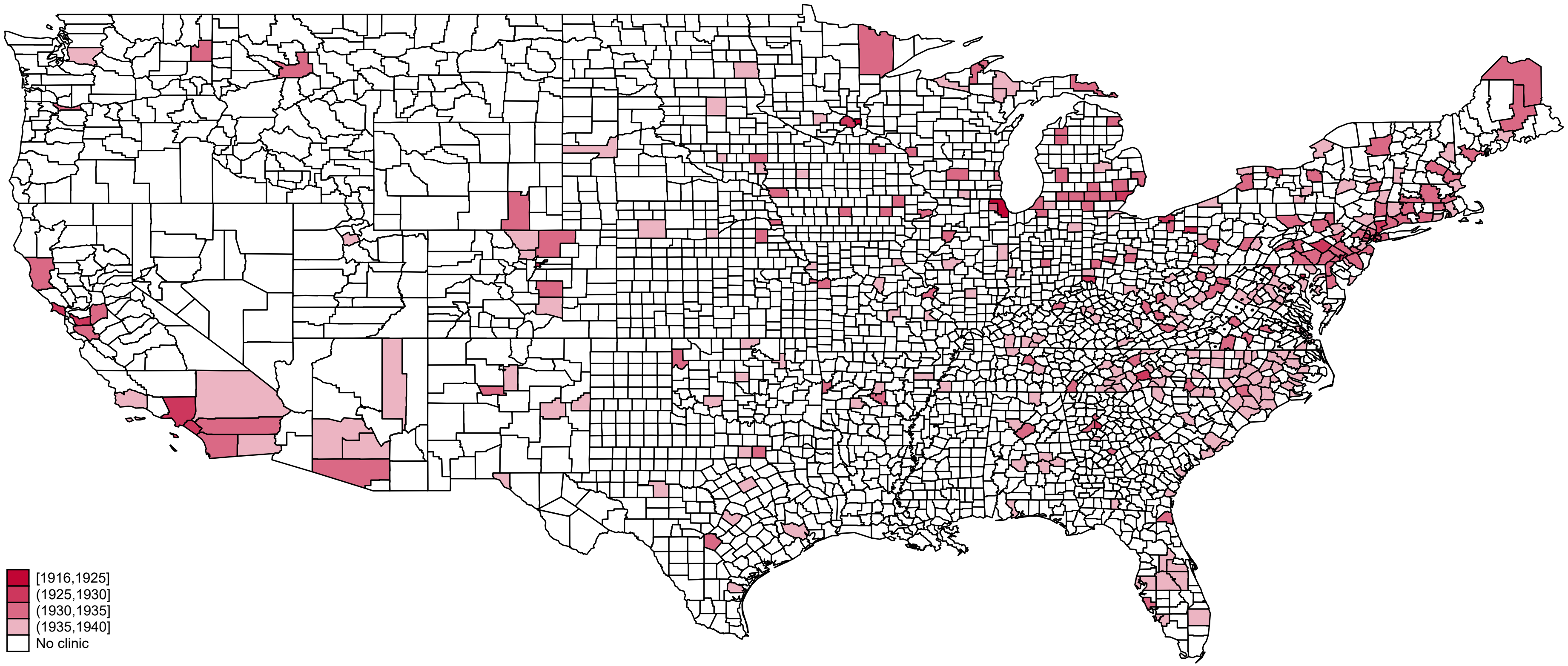
Notes: Data source: Birth control clinics statistics by Hajo (2010). The left panel shows the number of new birth control clinics by year. The right panel shows the cumulative number of new birth control clinics by year.

We depict the geographic dimension of the roll-out in Figure 3. The earlier a county was exposed to a birth control clinic, the darker it is colored. In principle, birth control clinics were first established in some of the bigger cities such as New York, Chicago, and Los Angeles. Later, they penetrated smaller cities and the countryside.[5] 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]

[5]Table A.1 in the Web Appendix shows these and other socio-economic characteristics over time in the groups of counties with and without exposure to a birth control clinic. The socio-economic variables are drawn from the Census, which we introduce in the next subsection.

[6]In particular, county borders might change if an existing county is divided in several new counties, if existing counties are merged to one new county, or if a county simply changes its geographical borders.

Figure 3 — 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.

It was not unusual that women attended a clinic in an adjacent county if there was no clinic in the county of residence. For example, Engelman (2011, p. 83) reports that women visiting the Brownsville clinic in Brooklyn came from all five boroughs of New York, from Long Island, from New Jersey and even from as far away as Philadelphia or New England. Therefore, in the empirical analysis, the variable measuring the existence of a birth control clinic or the exposure to it takes the value of one if a birth control clinic had opened either in a county itself or in one of the adjacent counties in or prior to the respective year.

3.2. Census data

To assess the impact of the exposure to a birth control clinic on the individual level, we use information on 15 to 39 years old women from the Population Censuses of 1920, 1930 and 1940. Harmonized full count data is available from the Integrated Public Use Microdata Series (IPUMS USA) (Ruggles *et al.*, 2020). We exclude the states of Alaska and Hawaii as well as the population living in military camps. Since we have information on the respondents' county of residence, we can merge the birth control clinic data to the Census data at the county level, again using the longitudinal template by Horan and Hargis (1995) with 1920 as the base year to work with consistent counties or county groups over time.

We limit our sample to White American and African American women as the remaining share of women is small (0.35%) and heterogeneous. They are either American natives or of Chinese, Japanese or a different Asian or Pacific origin. Since this group of women differs from White American and African American women in many dimensions, the effect of birth control clinics might also be markedly different. Moreover, we focus on marital fertility since birth control clinics did not serve to unmarried, childless women.[7]

We follow the historical literature and measure fertility using the child-woman ratio (see, e.g., Beach and Hanlon, forthcoming; Bleakley and Lange, 2009; Hacker, 2003). More precisely, we use the number of children below the age of five living in the household of each 15 to 39-year-old woman (Grimm, 2021). We 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. The number of children ever born is not available in the censuses of the years 1920 and 1930; it is available but vastly incomplete in the 1940 census. Relying on retrospective fertility in 1940 comes with two additional problems we would like to minimize, namely selective mortality and 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. To analyze potential effects of birth control clinics on females' labor market outcomes, we

[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.

use information on women's participation in the labor force (available in the 1920, 1930, and 1940 Census) and their employment status (available in the 1930 and 1940 Census).

The Census data further provides 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[8], her country of origin, 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. The census also classifies each location in terms of population size.[9] At the county level, we use information on the population's religious composition in 1916, i.e., the share of Catholics, Protestants, and followers of another religion, using data from IPUMS NHGIS (Manson *et al.*, 2021).

Since birth control clinics started to emerge in bigger cities we also conduct robustness checks using the sub-sample of 'big city counties' only. We define a county to be a 'big city county' if more than 50% of all women in a county lived in a city of more than 100,000 inhabitants in 1940. By fixing this category in the year 1940, we avoid that the group composition changes over time.

Table 1 shows descriptive statistics of the total sample, and separately for the 1920, 1930, and 1940 Census. In total, we observe more than 45 million women. The average woman in our sample is exposed to a birth control clinic for 2.75 years; this number increases from 0.246 in 1920 to 5.773 in 1940. The average number of children below the age of five is 0.63; it is highest in 1920 (0.747) and decreases to 0.541 in 1940. Women are on average 29 years old, 11% are African-American, and also 11% are foreign-born. The literacy rate is as high as 96%. 18% of all women report being in the labor force; the share of employed women is equally high but is only measured in 1930 and 1940. The county-level Protestant share in the total sample is 52%, while the Catholic share is 31%, and the share of followers of another religion is 17%. The share of women living in farm households is 22%. 57% of all women reside in urban areas, and 30% in big cities.

3.3. Administrative vital statistics

As a complementary data set, we use administrative county-level natality and mortality data put together 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. 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

[8] 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.

[9] The categories are places with less than 25,000, 25,000 to 50,000, 50,000 to 75,000, 75,000 to 100,000, 100,000 to 200,000, 200,000 to 300,000 etc.

Table 1 — Descriptive statistics: Census data

	1920	1930	1940	All years
Years of exposure to a BCC	0.246	1.550	5.773	2.745
	(0.959)	*(3.369)*	*(5.437)*	*(4.562)*
# childr. <5 in HH	0.747	0.626	0.541	0.629
	(0.876)	*(0.822)*	*(0.767)*	*(0.823)*
Age	29.333	29.518	29.544	29.474
	(5.910)	*(5.984)*	*(5.890)*	*(5.928)*
African American (=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
In labor force (=1)	0.131	0.176	0.215	0.178
Employed (=1)‡		0.166	0.199	0.183
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. † shares on county level. ‡ only available in 1930 and 1940.

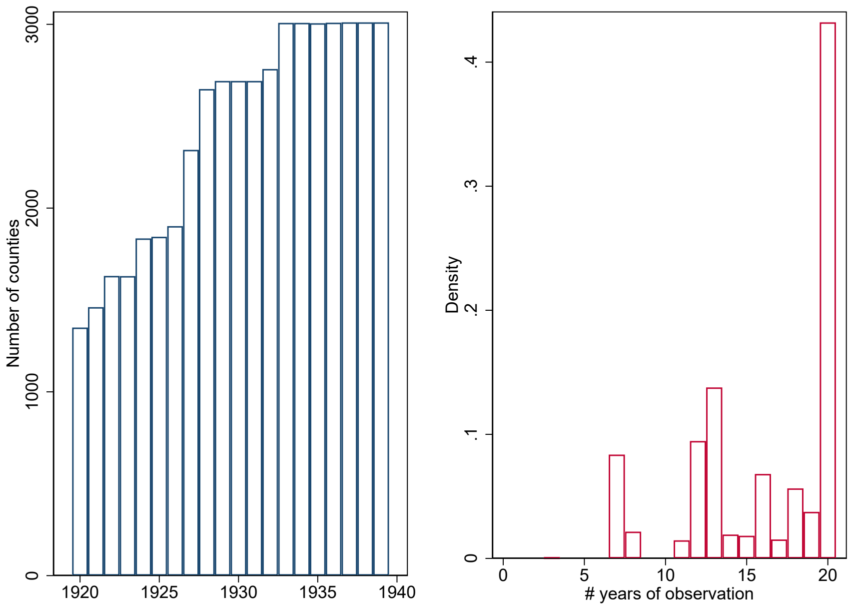
counties or county groups over time. This allows us to merge this data with the birth control clinic data at the county-year level.

For our period of observation from 1920 to 1940, Bailey *et al.* (2016) rely on four historical sources to collect the natality and mortality data. These are annual issues of Birth Statistics for the Registration Area of the United States (United States Bureau of the Census), annual issues of Birth, Stillbirth, and Infant Mortality Statistics for the Birth Registration Area of the United States (United States Bureau of the Census), annual issues of Birth, Stillbirth, and Infant Mortality Statistics for the Continental United States, the Territory of Hawaii, and the Virgin Islands (United States Bureau of the Census, 1922-1930; United States Public Health Service, 1931-1939) as well as annual issues of Vital Statistics of the United States: Part I (United States Bureau of the Census, 1937-1944; United States Public Health Service, 1945-1949). Moreover, Bailey *et al.* (2016) retrieve counties' full population counts for the years 1920, 1930, and 1940 from the Historical, Demographic, Economic, and Social Data: The United States, 1790-2002 (ICPSR 2896); they estimate population numbers for the years in between using a linear interpolation.

In our analysis, we use yearly information on the number of live births (exclusive of stillbirths), infant deaths (i.e., deaths of children below the age of one exclusive of stillbirths), and deaths by all ages. We subtract infant deaths from deaths by all ages to obtain a mortality measure unaffected by infant deaths. Births and deaths are assigned to the county of occurrence, and not to the county of residence in this period of observation. To account for changes in counties' population, we compute yearly 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.[10] For infant deaths, we use the number of births instead of the total population as the denominator.

Additionally, we use county level data from IPUMS NHGIS on the yearly number of stillbirths by the place of occurance in the period from 1922 to 1939 (Manson *et al.*, 2022). To obtain stillbirth rates, we divide this number by the sum of stillbirths and live births, where the latter variable comes from Bailey *et al.* (2016). Thus, the stillbirth rates take account of changes in fertility and do not automatically decrease as the number of pregnancies decreases.

FIGURE 4 — Overview of administrative data



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

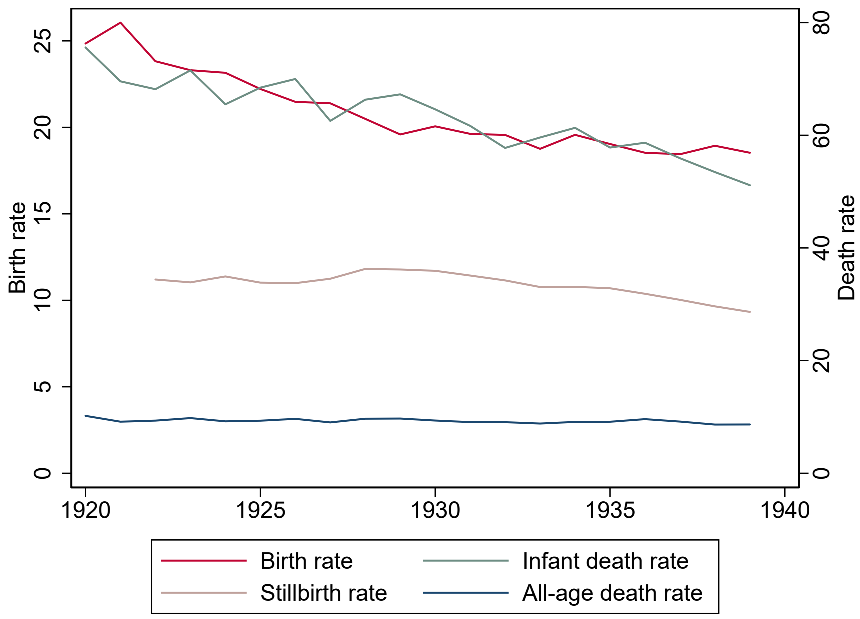
Figure 4 provides an overview of the number of observations available in the administrative data set. The left panel shows the number of counties in the data set by year. In 1920, we have information from 1,350 counties. This number steadily increases over the following years. From 1933 on, the data set includes observations from the universe of more than 3,000 counties. The right panel shows the distribution over the number of ob-

[10]Data on population by gender is not available before 1930. Therefore, we cannot compute fertility rates as the number of births by the number of women living in a county.

servations per county. For more than 40% of all counties, we have information from every single year in the period from 1920 to 1939. For 90% of all counties, we have observations from 11 years at least.

The development of fertility and mortality rates in our period of observation is depicted in Figure 5. 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 in 1920 to 51 per 1,000 births in 1939. Thus, we observe a 24% decrease in the birth rate and a 33% decrease in the infant death rate within twenty years. The number of stillbirths per 1,000 births (sum of live births and stillbirths) decreases from 34 in 1922 to 29 in 1939, which is a decline of 15%. All-age mortality (without infant deaths) is considerably smaller than infant mortality throughout the entire period of observation. It slightly decreases from 10 in 1920 to 9 in 1939.

FIGURE 5 — Birth and death rates in the administrative data



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

4. Evidence from Census data

4.1. Identification strategy

To identify the impact of a woman's exposure to a birth control clinic on the number of births, we estimate the following regression equation:

$$y_{ibcst} = \beta yexp^{BCC}_{ibcst} + x'_{ibcst}\gamma + \lambda_b \times \eta_s + u_{ibcst}, \qquad (1)$$

where y_{ibcst} is the number of children below the age of 5 in the household of woman i of birth cohort b, in county c, in state s, observed in census year t. $yexp^{BCC}_{ibcst}$ 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 does not only vary by county and year but even within a county-year cell across different ages of women. x'_{ibcst} is a matrix of control variables including age, age squared, literacy status, race, an indicator for being foreign born, an indicator for living in an urban area, an indicator for living in a big city, and an indicator for living in a farm household. It also includes the religious composition of the county a woman lives in. We include birth cohort fixed effects λ_b to capture general changes in fertility across birth cohorts. Each birth cohort groups women that are born within the same five-year period, starting with the period 1880–1885. By interacting the birth cohort fixed effects λ_b with state fixed effects η_s, we allow these cohort-specific changes in fertility to be heterogenous across states. This captures for example differences in anti-abortion and Comstock laws across states. In an extended specification, we additionally control for the county-level fertility rate in 1920 interacted with census year indicators. This allows us to capture county-level heterogeneity in fertility prior to the large expansion of birth control clinics, and account for differential county-level trends depending on the baseline fertility level. Standard errors u_{ibcst} are clustered at the birth cohort-county level because this is where the variation in the treatment comes from.

The coefficient of interest β measures the impact of one year of exposure to a birth control clinic under the assumption that, conditional on birth cohort by state fixed effects and the set of controls, the years of exposure to a birth control clinic are orthogonal to any other unobserved determinants of fertility. Holding fixed a birth cohort in a state, the variation we exploit comes from differences in the exposure to a birth clinic of a fixed birth cohort across counties of the same state.[11] Thus, the assumption can be reformulated in a way that it resembles the standard parallel trends assumption for difference-in-differences designs: we assume that women of the same birth cohort living in the same state follow the same fertility trend across counties in absence of the establishment of a birth control clinic.

To assess the validity of this key identifying assumption, we augment our analysis with an event study approach as laid out in the following equation:

$$y_{ibcst} = \sum_{\tau=-20}^{20} \beta_\tau exp^{BCC}_{ibcst,T+\tau} + x'_{ibcst}\gamma + \lambda_b \times \eta_s + u_{ibcst}. \tag{2}$$

$exp^{BCC}_{ibcst,T+\tau}$ takes the value one if a woman i of birth cohort b observed in county c, state s and year t is τ years prior to or after the first-time exposure. All other variables are

[11]To provide a better sense of the variation in exposure we exploit, Figure A.1 in the Web Appendix shows for each census year and for each cohort the mean exposure (a) measured by the share of women exposed and (b) measured by the mean years of exposure. As can be seen, exposure increases over time and is higher for younger than for older cohorts.

defined as in Equation 1. Thus, we compare women of the same birth cohort in the same state that are at different time distances from the event. In our preferred specification, we use intervals of three years instead of single years to increase statistical power. Thus, the first post-treatment interval runs from $\tau = 1$ to $\tau = 3$. Thereby, we also take into account that a prevented pregnancy needs at least nine months to "show up" in the data. In the regressions, we omit the category that captures years 16 or more prior to the treatment. Women from never-treated counties are subsumed under this category as well.[12]

If $\beta_{T+\tau}$ is zero for all $\tau < 0$, this indicates parallel fertility trends across counties in the pre-event period. Thus, this result would corroborate the validity of the key assumption of the empirical approach. Apart from allowing us to investigate pre-treatment trends, the event-study approach also allows us to trace treatment effects over the years after the event, i.e., 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, in an alternative model, we use county and time fixed effects instead of birth cohort by state fixed effects. Note that in this specification, we exploit variation within counties over time and across birth cohorts for identification. In another robustness test, we restrict the sample to big city counties to further enhance comparability of treatment and control group counties. Moreover, we check that the findings are confirmed in a Pseudo Poisson Maximum Likelihood regression since the number of children below the age of five is a count outcome variable. Furthermore, we follow Altonji *et al.* (2005) and Oster (2019) and compute bounds to our estimates based on assumptions about selection on unobservables. Finally, we use randomization inference techniques as an alternative way of deriving statistical inference.

4.2. Main results

We start our empirical analysis with a stripped down version of Equation 1, where we control for state-specific cohort fixed effects as well as age (and its square) and an urban area indicator. Controlling for women'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, we consider the urban area indicator as essential since it captures basic regional differences below the cohort-specific state fixed effects; in particular, this indicator accounts for the fact that birth control clinics were first established in urban areas before they penetrated more rural areas, while living conditions in urban and rural

[12] This is a natural choice since at some point in time after our period of observation, all women got access to birth control. We nevertheless test the robustness of our results with respect to this choice below.

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

	(I)	(II)	(III)	(IV)
Years of exposure to BCC	-0.004***	-0.004***	-0.003***	-0.003***
	(0.001)	(0.000)	(0.000)	(0.000)
State-specific cohort FE	yes	yes	yes	yes
Age and urban area controls	yes	yes	yes	yes
Socio-economic controls	no	yes	yes	yes
County level fertility 1920	no	no	yes	no
County level fertility 1920 × year FE	no	no	no	yes
R^2	0.072	0.084	0.088	0.088
Observations	45,104,489	45,104,489	45,104,489	45,104,489

*Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. The table shows OLS regressions. The dependant variable is the woman's number of children below the age of five living in the household. All regressions control for age and age squared as well as a variable indicating whether a houshold is in an urban area. 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 share of Catholics in a county. Standard errors are clustered at the birth cohort by county level. The detailed regression results are shown in the online appendix (Table A.2). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.*

areas were different at that time.

Column 1 of Table 2 shows a highly significant negative effect of exposure to a birth control clinic on the number of children below the age of five in a household. Adding the full set of individual level socio-economic controls in column 2 does not affect the estimate. The effect is not only statistically highly significant but also economically meaningful: being exposed to a birth control clinic throughout the entire fertile period from age 15 to 39 reduces fertility by about 15 percent ($\frac{25\times0.004}{0.629} \approx 15$). In column 3, we additionally control for county-level fertility measured in 1920, i.e, prior to the large expansion of birth control clinics, In column 4, we interact county level fertility 1920 with census year fixed effects to allow for differential county trends in fertility depending on the baseline fertility level. While the point coefficient slightly decreases, we still find sizeable and highly significant negative fertility effects of exposure to a birth control clinic. Exposure from age 15 to age 40 translates into a 12 percent ($\frac{25\times0.003}{0.629} \approx 12$) reduction in fertility.

In a next step, we move to the event-study specification of Equation 2 to inspect pre-treatment trends and post-treatment dynamics. Figure 6 shows the results of this event-study analysis, from which we obtain three insights. First, the analysis provides evidence for the validity of the common trends assumption. In particular, the coefficients for the time periods spanning years nine to one prior to the establishment of a birth control clinic are insignificant and close to zero. Although the two pre-treatment coefficients to the very left are significantly different from zero, the overall pattern of all pre-treatment coefficients does not suggest any gradual decrease in fertility before the establishment of a birth control clinic. Second, we observe an immediate significant reduction in fertility once the birth control clinic is established. Third, this negative effect is persistent and even tends to grow larger over the post-treatment years at least until year sixteen. Because only few women are exposed to a birth control clinic for more than sixteen years, the confidence intervals for these long-term coefficients are rather large.[13]

As fertility declines with exposure to birth control clinics, we might expect positive knock-on effects on female labor supply. If birth control clinics increase the spacing of births (and thus reduce the total number of births), women might be more likely to re-enter the labor market between births. We empirically test this hypothesis by using an indicator for a woman's labor force participation (available in the Census years 1920, 1930, and 1940) and a woman's employment status (available in the Census years 1930 and 1940) as the dependent variables of Equation 2. Table A.3 and Figure A.3 in the Appendix provide suggestive evidence for a positive effect of exposure to birth control clinics on female labor force participation and employment. Yet, since ambiguous pre-trends partly blur the overall picture, we are cautious and do not over-interpret the effects on female labor market participation. Moreover, the beginning of the Great Depression in 1929 had a substantial impact on labor demand and might have reduced women's opportunities to enter the labor market.

4.3. Robustness tests

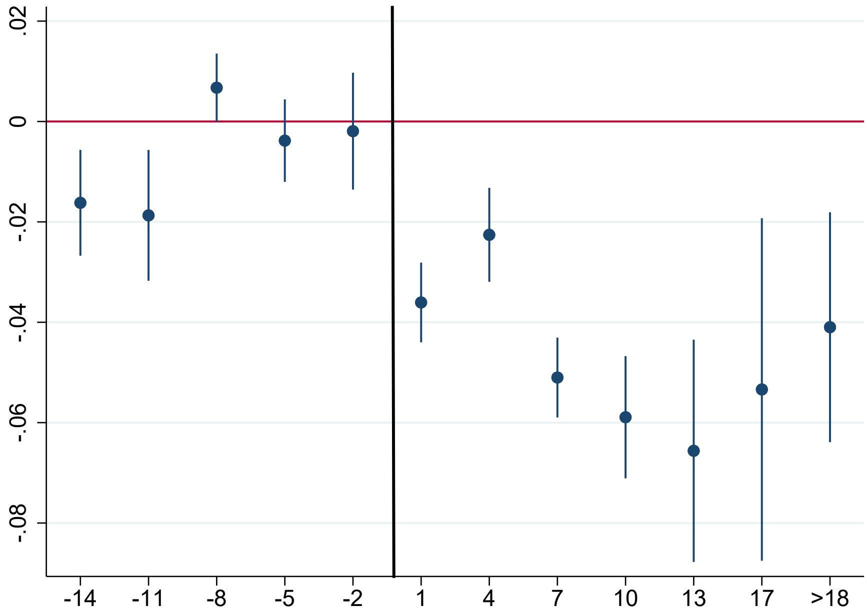
To test the robustness of our findings, we now include county and year fixed effects instead of birth cohort-specific state effects in the model. The inclusion of county fixed effects allows us to capture time-invariant regional heterogeneity below the state level, while the year fixed effects take account of general time trends in fertility. Column 1 of Table 3 shows that our results remain virtually unchanged in this alternative specification.

As another way of taking into account regional heterogeneity, we re-run the regression along the lines of Equation 1 on the subsample of counties with big cities. Since the roll-out of birth control clinics started in big cities and only later penetrated rural areas, this procedure should also help making treatment and control groups more comparable to each other. From column 2 of Table 3 we can see that our findings are confirmed on the subsample of big city counties. In big city counties, the effect is even larger suggesting that birth control might have met a higher demand in big cities than in smaller cities and rural areas.

Since the number of children below the age of five in a household is a count outcome variable, we re-estimate the model using Pseudo Poisson Maximum Likelihood regressions. Column 3 of Table 3 confirms that exposure to a birth control clinic significantly reduces

[13] We obtain virtually identical results if we restrict the sample to women from ever-treated counties (see Figure A.2 in the Web Appendix) instead of subsuming women from never-treated counties under the omitted category capturing years 16 and more prior to the treatment.

Figure 6 — Event-study plot 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 children below the age of five in a household. Treatment effects are estimated along the lines of equation 2. The whiskers mark the 95 percent confidence band.

fertility. Indeed, the effects turn out to be even larger in the Pseudo Poisson Maximum Likelihood model than in the OLS model. Every year of exposure reduces fertility by 1.3 percent, as compared to 0.6 percent in the OLS model.

Following the methodology developed by Altonji *et al.* (2005) and further elaborated by Oster (2019), we test how robust our coefficient of interest is to selection on unobservables. This test is based on the idea that the shift in the R^2 in relation to the shift of the coefficient of interest resulting from the inclusion of observed controls carries information on the shift of the coefficient of interest in case relevant unobservables could be included. This test requires an assumption by how much the R^2 could increase in this latter case. We follow Oster (2019) who suggests, based on illustrative examples, a value of 1.3. We compare the stripped down model (column 1 of Table 3) to the model including the full set of observable socio-economic characteristics (column 2 of Table 3). Column 4 of Table 3 shows that the bias-adjusted coefficient of interest is still negative and comparable in size to coefficients from our main specifications in Table 3. The effect would be eliminated if the ratio of the selection on unobservables to selection on observed socio-economic characteristics was 4.090.

Finally, we conduct randomization inference tests (Rosenbaum, 2002; Cohen and Dupas, 2010) based on 100 permutations to derive statistical inference for the birth cohort by state as well as the county fixed effects model. Each permutation allocates the en-

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

	County FE (I)	Big cities (II)	Poisson (III)	Oster bounds (IV)
Years of exposure to BCC	-0.003***	-0.005***	-0.013***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
bias-adjusted coeff.				*-0.003*
δ eliminating effect				*4.090*
State-specific cohort FE	no	yes	yes	yes
Age and urban area controls	yes	yes	yes	yes
Socio-economic controls	yes	yes	yes	yes
County FE	yes	no	no	no
Year FE	yes	no	no	yes
R^2	0.087	0.066	0.045	0.084
Observations	45,104,489	15,156,837	45,104,489	45,104,489

*Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940 and NHGIS. The table shows OLS regressions. The dependant variable is the woman's number of children below the age of five living in the household. All regressions control for age and age squared as well as a variable indicating whether a houshold is in an urban area. 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 share of Catholics in a county. Standard errors are clustered at the birth cohort by county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.*

tire set of exposure start dates randomly to counties. The distribution of coefficients obtained on the permuted samples includes zero. The point estimate obtained from the actual distribution of the birth control clinics is clearly outside of that distribution (see Figure A.4 in the Online Appendix).[14] Thus, we conclude that it is unlikely that the negative coefficients are just the result of statistical uncertainty.

5. Evidence from administrative data

Although the Census data provide a sound basis for addressing our research question, several issues remain. First, we cannot directly observe yearly births but need to rely on the respondents' information about the number of children below the age of five in a household. Second, since the Census data are collected every ten years, we can only use information from three points in time, namely 1920, 1930, and 1940. Third, we cannot observe deaths although fetal deaths, infant deaths and maternal deaths are interesting outcome variables in our setting. To address these issues, we now complement the empirical analysis with a second approach that uses yearly administrative data on the universe

[14]The fact that the distribution of coefficients obtained on the permuted samples for the birth cohort by state model (left panel) is not centered around zero is due to the fact that states with more counties are more likely to have clinics randomly allocated to them and states with more counties were more urbanized and experienced a faster and earlier fertility decline than those counties that are in states with fewer counties.

of live births, stillbirths and deaths at the county level.

5.1. Plotting raw data

We start the empirical analysis of the administrative data with simple descriptive plots. For the group of *ever-treated* counties, we normalize the year of the opening of the first birth control clinic in a county or an adjacent county to zero. For the group of *never-treated* counties, we randomly assign placebo openings of birth control clinics. To this end, we first take the sample of *ever-treated* counties and compute for every single year from 1916 tp 1939, which share of counties was affected by the opening of the first birth control clinic. We take the resulting distribution of probabilities as the basis for our random assignment of placebo opening to *never-treated* counties. Then, we normalize the year of the placebo opening for *never-treated* counties to zero.

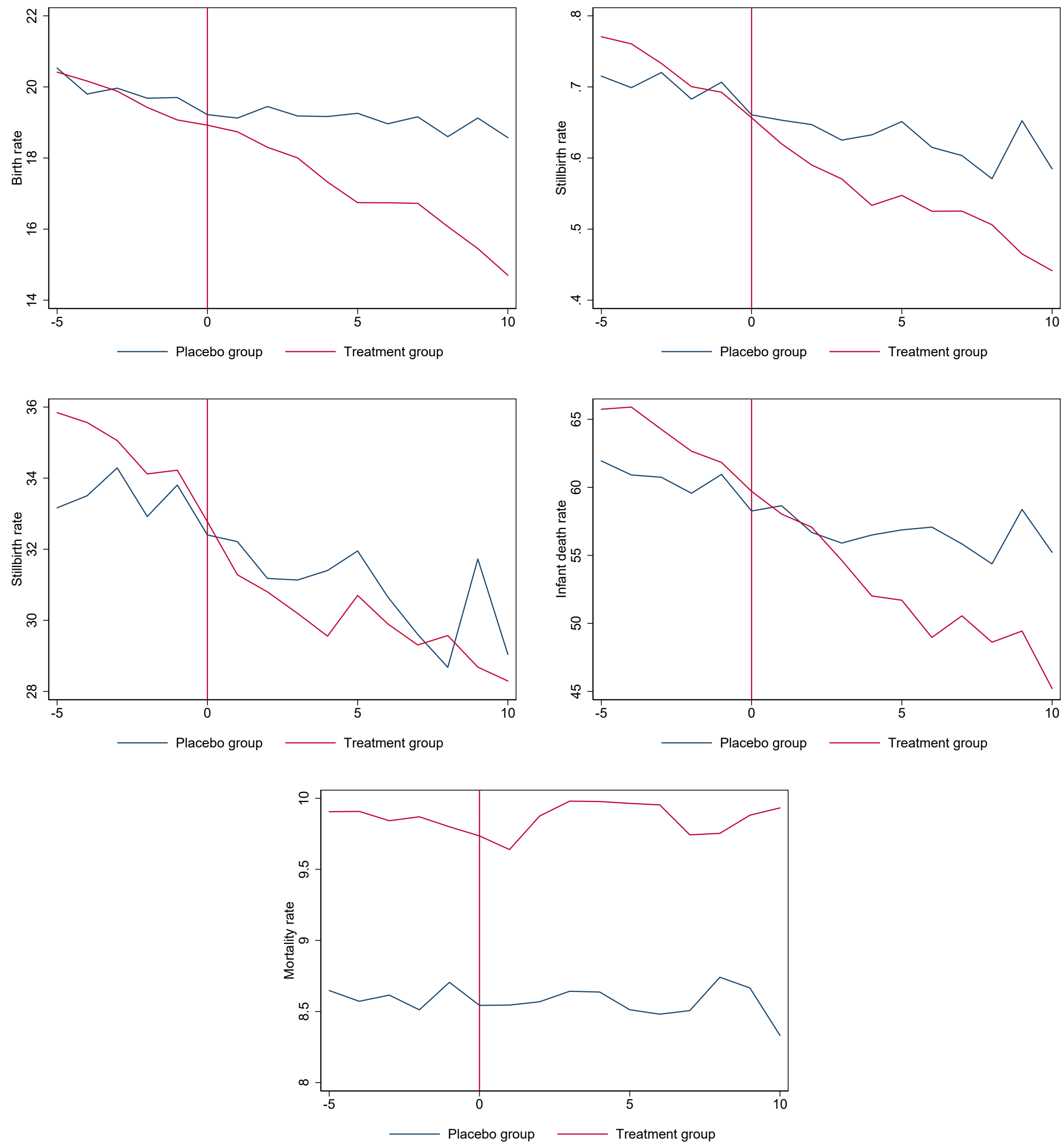
Figure 7 plots the birth rate, the stillbirth rate, the infant death rate, and the all-age mortality rate for the treatment and the placebo treatment group against time from the (placebo) opening of a birth control clinic. For the birth rate (upper left graph), we observe a stronger decrease for the treatment group than for the placebo treatment group after the (placebo) opening of a birth control clinic. The same is true for the number of stillbirths per 1,000 population (upper right panel). If we use the number of stillbirths per 1,000 births, the pattern is somewhat less clear (middle left figure). Infant mortality seems to decrease more strongly for the treatment group than for the control group (middle right figure). The results for the all-age mortality rate without infant deaths (bottom figure) are more ambiguous. Thus, these first descriptive results are in line with the results from the Census data in suggesting a negative effect of birth control clinics on fertility. Moreover, they suggest that birth control clinics might have had a negative effect on the stillbirth and the infant death rate, while the pattern is less clear for the all-age mortality rate (without infant deaths).

For the remaining analysis, we create a balanced panel of counties over the period from 1925 to 1939. As explained in Section 3, the number of counties available in the administrative data 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.[15] Using the balanced panel, we match untreated counties to treated counties using nearest neighbor matching techniques and drop unmatched counties from the data set.[16] Proceeding like this, we hold the composition of the sample constant over

[15] In line with Sun and Abraham (2021), we drop counties treated already prior to the first year we observed in this balanced panel.

[16] Table A.4 in the Appendix 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.

Figure 7 — Raw difference-in-differences plots for fertility and mortality



Notes: Data source: U.S. Vital Statistics. The figure shows the evolution of fertility rates, stillbirth rates, infant death rates, and all-age mortality rates for the group of ever-treated counties and the group of never-treated counties. For never-treated counties, we randomly assign placebo opening years of birth control clinics. The year of an actual or a placebo opening are normalized to $t = 0$.

time and ensure that treatment and control counties are comparable in observed baseline characteristcs. Table A.5 in the Appendix shows that this main sample hardly differs from the unbalanced sample of all counties from 1920 to 1939 with respect to the means of the fertility and mortality outcome variables.

5.2. 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 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). From Figure A.5 in the Appendix it becomes evident that this is a problem also in our setting. The relative period coefficients are weighted combinations of own period effects and other period effects, and the weights are sometimes negative. As a result, the event-study coefficients estimated in a dynamic two-way fixed effects model are hard to interpret.

We avoid these pitfalls by estimating an event-study applying Sun and Abraham (2021)'s interaction weighted estimator. The estimator is constructed in three steps. First, we run a linear two-way fixed effects specification and interact indicators for relative event period l with cohort indicators, where the cohort e is defined as the group of counties i whose treatment starts in the same year. The event, i.e., the start of the absorbing treatment, is determined by the year in which the first birth control clinic in a county or an adjacent county was established; this year constitutes the omitted category. Cohorts that are always treated are excluded from the estimation, while the *never-treated* cohort forms the control group C.

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

Second, we estimate the cohort shares in each relative time period 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 shares 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. Figure A.6 in the Appendix provides an overview of the relative time specific weights. The interaction weighted estimator depicts the causal effect of birth control clinics under the assumptions of parallel trends in absence of the treatment and no anticipation of the treatment.

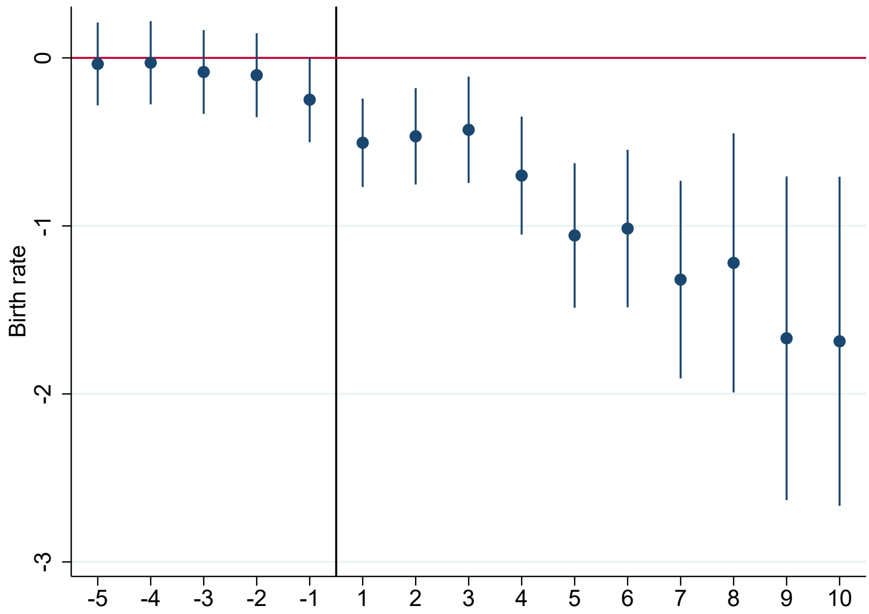
We provide evidence for the validity of the parallel trends assumption by inspecting pre-trends. In particular, if the estimates of the leads in the event-study turn out to be zero, trends in the periods prior to the opening of a birth control clinic are parallel, which makes it plausible that trends 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. Since we have yearly data, we can not only compare the general trend in the leads but also, in particular, the difference between the event year and the year immediately prior to the event. This should give us some reliable insights into the empirical relevance of anticipating behavior.

5.3. Effects on fertility

Figure 8 shows the event study estimates for the impact of birth control clinics on fertility. All pre-treatment coefficients are close to zero and insignificant, which suggests that in the years before the opening of a birth control clinic, treatment group counties follow a similar birth rate trend as never-treated counties. This finding corroborates the validity of the key identifying assumptions. 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 even tends to grow in the course of ten years after the opening of the birth control clinic. Averaged over the period of ten years after opening, birth control clinics reduce the local annual birth rate by 1.007 births, which amounts to 5.2 percent of the sample mean or 16 percent of the standard deviation.

Comparing these estimates based on administrative data to the Census data estimates, we find that the effect sizes are indeed similar. In particular, we compare the effect of 25 years of exposure to a birth control clinic in the Census data with the effect of availability of a birth control clinic in the administrative data. Thus, in the Census data, we look at cohort exposure, while in the register data, we look at the exposure of a cross-section of women of fertile age (roughly 25 age-years) in a given year. Based on the Census data, we find that 25 years of exposure to a birth control clinic reduces the number of children below five in a household by 11 percent of a standard deviation. Based on the administrative data, we find that exposure to a birth control clinic reduces the birth rate by 16 percent of a standard deviation. Yet, note that comparability might suffer from the fact that both the time period and the set of counties we use differs between the Census data and the administrative data estimations. While we can draw on data from all counties in the period from 1920 to 1940 in the Census data, the administrative data restrict us to a panel of counties that consistently report vital statistics from 1925 to 1939, from which we additionally drop unmatched counties. This reduces the number of distinct counties from 3,011 to 1,275.

Figure 8 — 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. Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). We bin ten or more years after the treatment into a single indicator. The control group consists of never-treated counties. The whiskers mark the 95 percent confidence band. The sample is restricted to a balanced panel of counties for the years 1925 to 1939.

5.4. 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. Conde-Agudelo *et al.* (2006) provide a systematic literature review of the role of birth spacing for perinatal health. This review finds that interpregnancy intervals shorter than 18 months or longer than 59 months are negatively associated with a range of perinatal outcomes. More recently, Gupta *et al.* (2019) show in a case control study that shorter intervals between pregnancies increase the risk of stillbirths. This effect is not confounded by the fact that women with a prior pregnancy loss are more likely than women with a live birth to shorter spacing between pregnancies; rather, short spacing constitutes an independent risk factor. Molitoris (2017) exploit within-family variation in the association between preceding birth intervals and infant mortality risks using early 20th century register data from Sweden. He finds evidence for a negative impact of short interpregnancy intervals on infant mortality. Conde-Agudelo *et al.* (2012) review causal 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 the most important channels.

Figure 9 presents evidence on the impact of birth control clinics on stillbirths. In the upper panel, we use the number of stillbirths per 1,000 population, while in the lower panel, we use the number of stillbirths per 1,000 births (sum of stillbirths and live births) as the outcome variable of the regression. For both outcome variables, all pre-treatment coefficients are close to zero and 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. The fact that we do not only see this decrease for the number of stillbirths per 1,000 population but also for the number of stillbirths per 1,000 births suggests that the decrease of the number of stillbirths is not just a mechanical effect arising from the decrease in pregnancies. Rather, the results suggest that birth control clinics avert particularly high risk pregnancies by increasing the spacing of births. Averaged over ten years after opening, birth control clinics reduce the number of stillbirths per 1,000 births by 1.482, which amounts to 4.5 percent of the sample mean or 9.6 percent of a standard deviation.

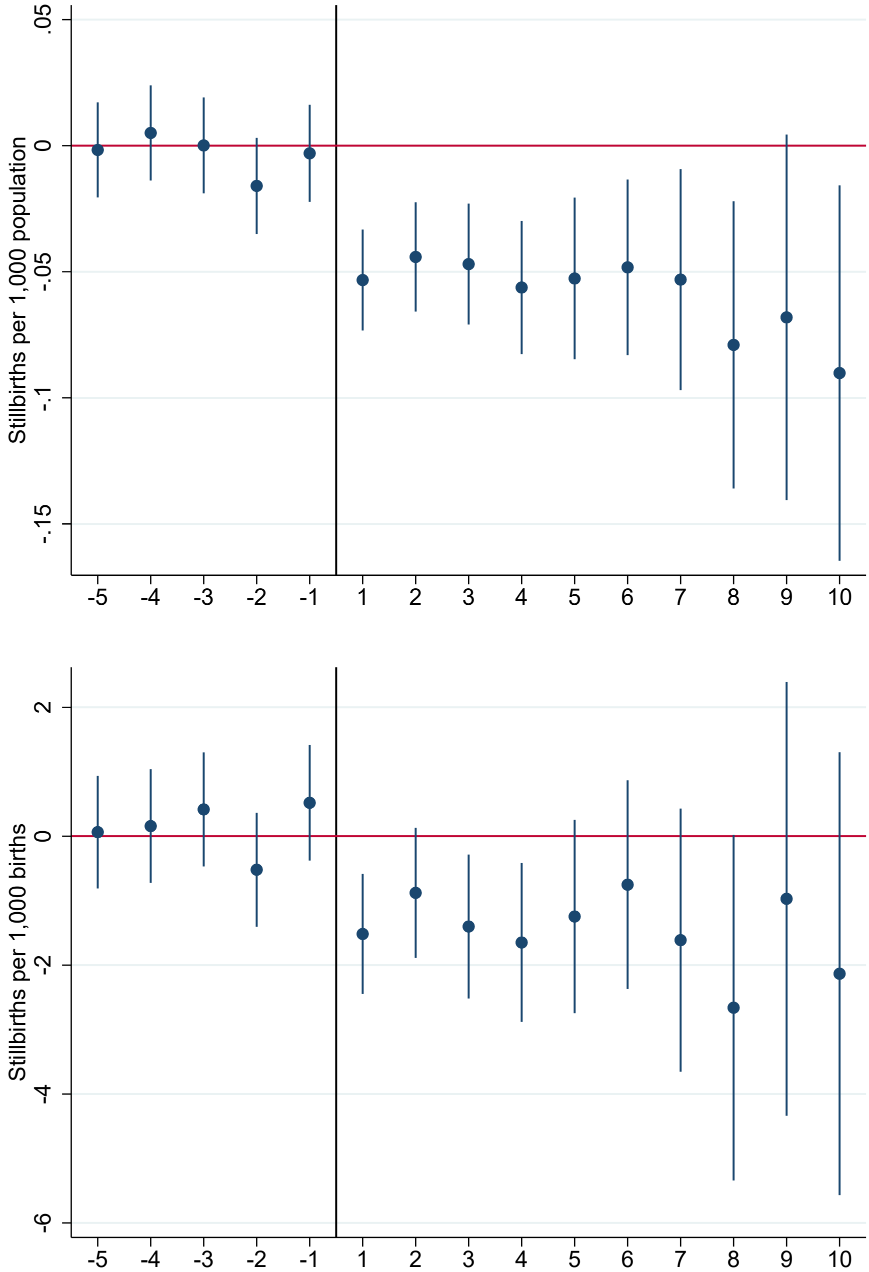
In a next step, we investigate the impact of birth control clinics on infant mortality. As we can see in Figure 10, all five pre-treatment coefficients are insignificant and virtually zero, which corroborates the validity of the common trends and the no-anticipation assumption. After the opening of a birth control clinic, the infant death rate significantly drops; in the following years, the 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 do not just reduce the number of births but indeed improve the average health of the babies born. Again, this means that birth control clinics particularly avert the type of births that are characterized by increased health risks. Averaged over the period of ten years after opening, a birth control clinic reduces the local infant death rate by 4.206 deaths, which amounts to 7.0 percent of the sample mean or 19.6 percent of a standard deviation.

5.5. Effects on maternal mortality

Since birth control clinics reduce the number of births, they might also reduce the number of puerperal deaths. Moreover, birth control clinics might reduce maternal deaths via an increased birth spacing as such. In their review article, Conde-Agudelo *et al.* (2012) report increasing evidence that an incomplete healing of the uterine scar from a previous cesarean delivery is a major risk factor for maternal health in case of short interpregnancy intervals. Short birth spacing might also result in maternal nutritional depletion, which constitutes another risk factor, although the empirical evidence for this channel is still inconclusive. Furthermore, access to contraceptives via birth control clinics might have reduced unprofessional abortions and thereby reduced health risks.

Figure 11 shows event-study estimates for the impact of birth control clinics on all-age mortality. All pre-treatment coefficients are insignificant and close to zero, which again corroborates the validity of the identifying assumptions. After the opening of the birth

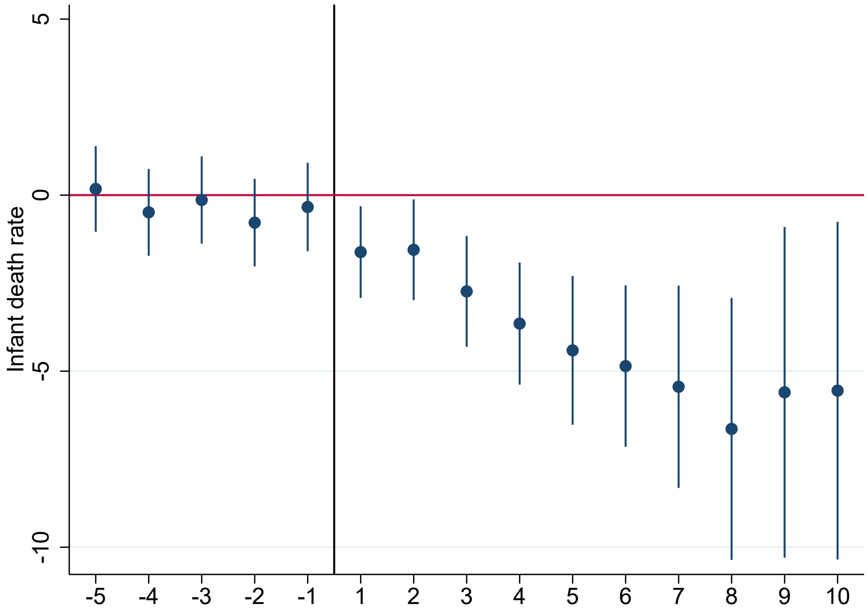
Figure 9 — The impact of birth control clinics on stillbirths (Sun and Abraham, 2021)



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 (upper panel) and stillbirths per 1,000 births (lower panel). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). We bin ten or more years after the treatment into a single indicator. The control group consists of never-treated counties. The whiskers mark the 95 percent confidence band. The sample is restricted to a balanced panel of counties for the years 1925 to 1939.

control clinic, we do not find any evidence for significant effects on all-age mortality. One reason for this result might be that we cannot focus on the mortality rate of females of reproductive age (or the maternal mortality rate) but need to rely on the total all-age mortality rate due to a lack of gender and age specific mortality data in this period.

Figure 10 — The impact of birth control clinics on infant mortality (Sun and Abraham, 2021)



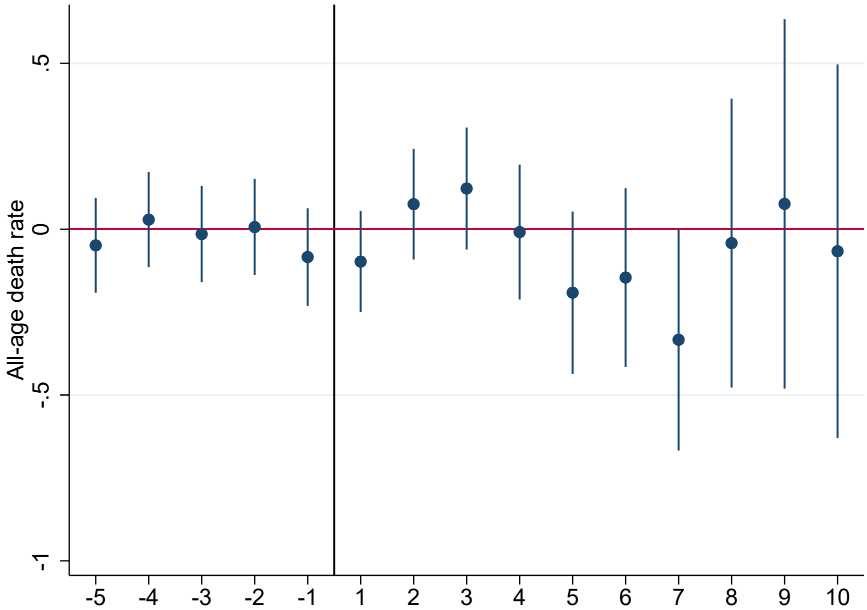
Notes: Data source: U.S. Vital Statistics. The figure shows the dynamic impact of the establishment of the first birth control clinic in a county (or an adjacent county) on the infant death rate. Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). We bin ten or more years after the treatment into a single indicator. The control group consists of never-treated counties. The whiskers mark the 95 percent confidence band. The sample is restricted to a balanced panel of counties for the years 1925 to 1939.

To address this issue, we now exploit administrative city-level data on causes of death. The *U.S. Department of Commerce* published cause-specific deaths for the 100 biggest U.S. cities from the 1930s onward for selected years (later on a yearly basis). Although this data does not report deaths by gender either, it allows us to identify the number of puerperal deaths, which are most closely associated with potential health effects from birth control clinics. To account for the size of the underlying population, we divide the number of puerperal deaths by the female population aged 15 to 44, i.e. the population susceptible to give birth. This data is available from Manson *et al.* (2021) for the Census years 1930 and 1940, and interpolated for the years in between. For our analysis, we use the data for the years 1931 and 1939 and restrict the sample to cities that consistently report the number of puerperal deaths as well as the number of cancer and heart diseases deaths that we use as placebo outcomes.

We run a double difference model to estimate the effects of birth control clinics on the puerperal death rate.[17] In particular, we regress a city's puerperal death rate on group fixed effects (ever-treated vs. never-treated), year fixed effects (1939 vs. 1931), and the interaction of the group and the 1939 fixed effect. The coefficient of the interaction effect

[17]Thereby, we avoid problems that might arise in a staggered treatment design based on observations from few years only where, additionally, the interval between observed years varies.

Figure 11 — The impact of birth control clinics on all-age 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 all-age mortality rate (without infant deaths). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). We bin ten or more years after the treatment into a single indicator. The control group consists of never-treated counties. The whiskers mark the 95 percent confidence band. The sample is restricted to a balanced panel of counties for the years 1925 to 1939.

tells us by how much the change in the puerperal death rates differed between those cities that had opened a birth control clinic by 1939 and those cities that had not opened a birth control clinic by 1939.[18]

We start with an OLS regression and check the robustness in a Pseudo Poisson Maximum Likelihood model due to the small number of disease-specific cases. Table 4 provides evidence for negative effects of birth control clinics on puerperal deaths. The effects are imprecisely measured but they are economically meaningful. The point estimate from column 1 suggests that cities with a birth control clinic could reduce the puerperal death rate by about 44 percent more than other cities. The effect is in the order of 32 percent if alternatively the Poisson model is used in column 2.

[18]If we drop cities that opened a birth control clinic prior to 1931, the number of observations further decreases to 74 cities and the difference-in-differences model yields more ambiguous results.

Table 4 — Mortality by cause 1931 and 1939, 100 biggest cities

	(I)	(II)	(III)	(IV)	(V)	(VI)
	Puerperal		Cancer		Heart diseases	
	OLS	Poisson Pseudo-ML	OLS	Poisson Pseudo-ML	OLS	Poisson Pseudo-ML
Birth control clinic	-0.152*	-0.320**	-0.020	-0.018	0.159	0.063
	(0.086)	(0.159)	(0.067)	(0.054)	(0.151)	(0.059)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Ever treated fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep var (cases per 1000)	0.345	0.345	1.300	1.300	2.778	2.778
Observations	178	178	178	178	178	178
Cities	89	89	89	89	89	89
(Pseudo) R-squared	0.162	0.072	0.087	0.003	0.251	0.018

Notes: Data sources: U.S. Vital Statistics for the years 1931 and 1939. Column I and II considers death cases per female population 15-44 years. Data on mortality due to puerperal disease is missing for 5 cities. Columns III-VI consider cases per 1,000 population. Standard errors are clustered at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To check the validity of this approach, we run placebo tests using cancer and heart disease death rates as placebo outcomes. These causes of death are unlikely to be affected by birth control and hygiene measures that were propagated by birth control clinics. We compute cause-specific death rates by dividing the number of deaths due to heart diseases and due to cancer by the total population of a city. Columns 3 and 4 of Table 4 show that we do not find any statistically significant effects for the placebo outcomes. Thus, these placebo tests corroborate the validity of the empirical approach. Taken together, these results provide evidence that birth control clinics had a negative effect on maternal deaths due to fewer puerperal deaths.

5.6. Robustness tests

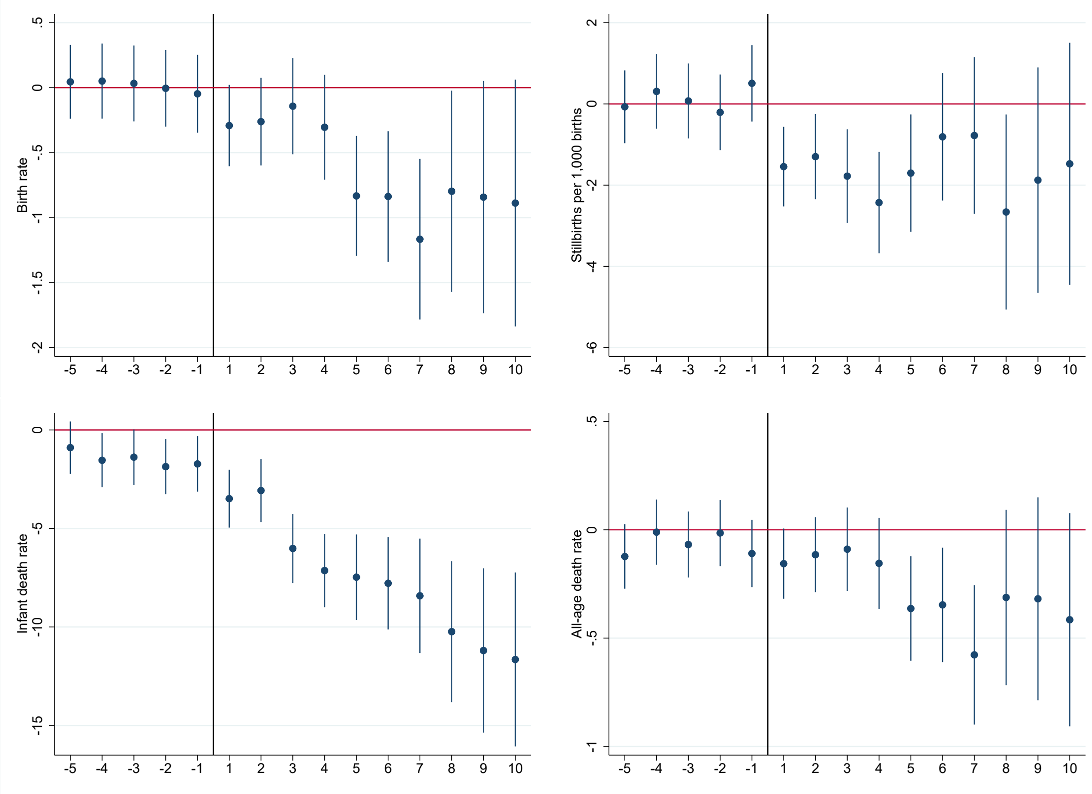
As a first robustness test, we check whether the results are confirmed if we use a 20 years balanced panel instead of the 15 years balanced panel. On the one hand, this reduces the number of counties in our data set. In particular, counties that enter the vitality statistics between 1921 and 1925 are not included in the 20 years balanced panel from 1920 to 1939, while they are included in the 15 years balanced panel from 1925 to 1939. On the other hand, the 20 years balanced panel allows us to draw on more yearly data points for the set of counties that are included in the panel. 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 18,300 observations from 915 counties.[19] Figure 12 reports the event study estimates on this alternative sample and confirms all previous findings. The effects of birth control clinics on stillbirths and infant mortality are even more pronounced than in the 15 years balanced panel. For all-age mortality, the event study now yields consistent negative effects that turn significant in three of the ten post-treatment periods.

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).[20] Figure 13 presents the results of this alternative specification. The 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. In contrast, the negative effect of birth control clinics on stillbirths and the infant mortality rate is clearly confirmed. Also for for all-age mortality, the pattern is similar as in the main specification.

[19] Table A.5 in the Appendix shows that this sample hardly differs from the 15 years balanced panel of the main specification.

[20] Table A.5 in the Appendix shows that this sample hardly differs from the 15 years balanced panel of the main specification.

Figure 12 — The impact of birth control clinics in a 20 year panel (Sun and Abraham, 2021)



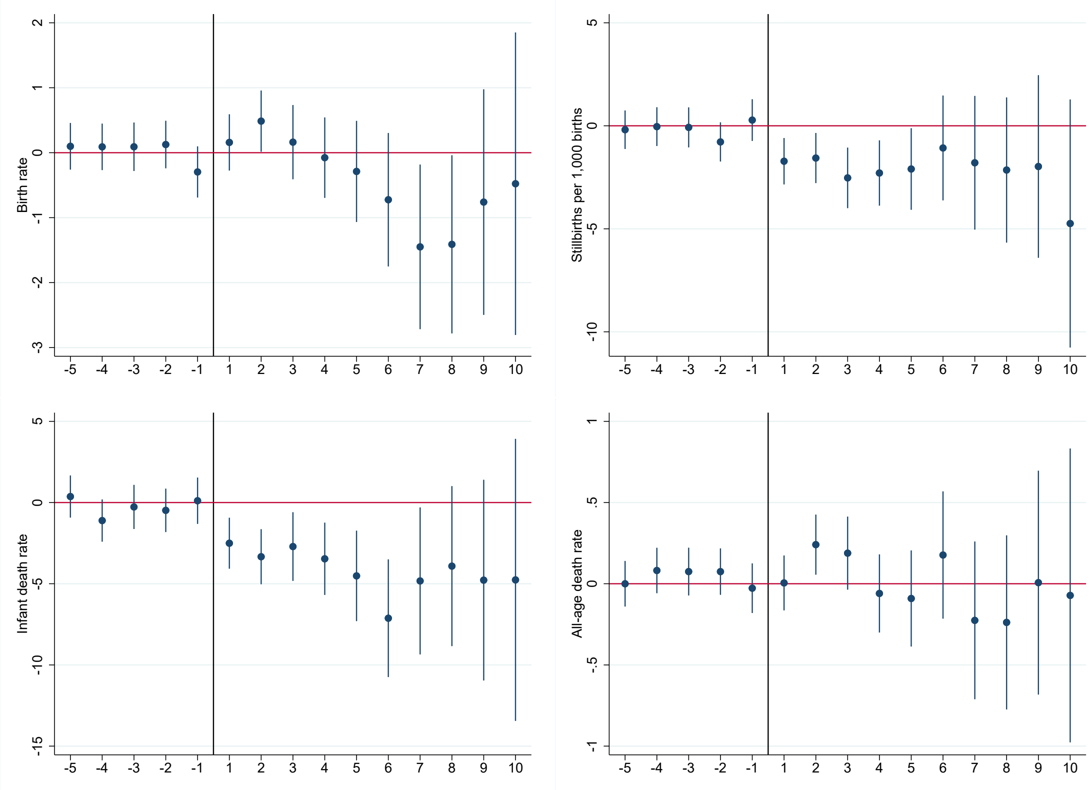
Notes: Data source: U.S. Vital Statistics. The figure shows the dynamic impact of the establishment of the first birth control clinic in a county (or an adjacent county) on the birth rate (upper left panel), the stillbirth rate (upper right panel), the infant death rate (lower left panel) and the all-age death rate (lower right panel). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). We bin ten or more years after the treatment into a single indicator. The control group consists of never-treated counties. The whiskers mark the 95 percent confidence band. The sample is restricted to a balanced panel of counties for the years 1920 to 1939.

6. Conclusions

Margaret Sanger was the leading figure of the early 20th century U.S. birth control movement. In opposition to the incumbent conservative religious and political leaders, she devoted her life to liberalize the use of contraceptives and advice women on sexual matters and hygiene. Her ultimate aim was to improve women's and children's health and to relief poverty among working class women. With this motivation, Sanger opened the United States' first birth control clinic in New York in 1916. Hundreds of women visited the clinic before it was shut-down by the police after ten days. Yet, this incidence as well as repeated criminal accusations did not stop Sanger's movement. The country-wide roll-out of Sanger's birth control clinics started in 1923 and gained speed in the 1930s. By 1940, more than 600 birth control clinics were established throughout the U.S.

To assess the causal impact of Sanger's birth control clinics on U.S. fertility and mortality, we combine four data sets. At the heart of our analysis are newly digitized data on the roll-out of birth control clinics. We merge this data with full-count Census

Figure 13 — 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 (upper left panel), the stillbirth rate (upper right panel), the infant death rate (lower left panel) and the all-age death rate (lower right panel). Treatment effects are derived using the interaction weighted estimator by Sun and Abraham (2021). We bin ten or more years after the treatment into a single indicator. The control group consists of the counties that are treated in 1938 and 1939. The whiskers mark the 95 percent confidence band. The sample is restricted to a balanced panel of counties for the years 1925 to 1937.

data of 1920, 1930 and 1940, yearly administrative vital statistics at the county level, and administrative data on causes of deaths at the city level. Using the Census data, we estimate the causal effect of birth control clinics on the number of children below five in a household. To this end, we draw on within birth cohort by state variation in the length of exposure to clinics while controlling for women's age and a set of additional individual level and county level covariates. Using the yearly administrative data at the county level, we estimate the causal impact of birth control clinics on the birth rate, the stillbirth rate, the infant mortality rate, and the all-age mortality rate. To this end, we exploit the staggered roll-out of birth control clinics and estimate event study models employing Sun and Abraham (2021)'s interaction weighted estimator. Finally, we use city-level data on causes of death to better capture the type of (maternal) deaths most likely affected by birth control clinics, namely puerperal deaths.

Our estimates using the Census data show that being exposed to a birth control clinic for the entire fertile period from age 15 to 39 reduces fertility by 12 to 15%. Administrative data confirm the negative impact of birth control clinics on fertility. Moreover, we find evidence that birth control clinics reduce stillbirths and infant mortality. Further, birth control clinics reduce maternal deaths by reducing the number of puerperal deaths. These findings suggest that birth control clinics increase birth spacing and thus particularly avert births that are characterized by high health risks. We also find some suggestive evidence for positive effects on female labor supply. Parallel pre-treatment trends provide evidence for the validity of the empirical approaches. Various specification checks including placebo outcomes, alternative control groups, and subsample analyses confirm the findings.

Overall, we interpret these findings as the consequence of a supply side policy that relaxed constraints on the demand side, i.e., in a Beckerian framework (Becker, 1981) birth control clinics reduced the cost of having fewer children. Thus, we provide new evidence on the determinants of the U.S. demographic transition at the beginning of the 20th century.

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

Altonji, J. G., Elder, T. E. and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, **113** (1), 151–184.

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.

—, Guldi, M. and Hershbein, B. J. (2014). Is there a case for a second demographic transition? Three distinctive features of the post-1960 U.S. fertility decline. In *Human Capital in History: The American Record*, NBER Chapters, National Bureau of Economic Research, pp. 273–312.

— 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.

Beach, B. and Hanlon, W. (forthcoming). Culture and the historical fertility transition. *Review of Economics Studies*.

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

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.

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.

Cohen, J. and Dupas, P. (2010). Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *The Quarterly Journal of Economics*, **125** (1), 1–45.

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.

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.

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. (2007). 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.

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

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.

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.

— 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.

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.

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.

Manson, S., Schroeder, J., Van Riper, D., Kugler, T. and Ruggles, S. (2021). *IPUMS National Historical Geographic Information System: Version 16.0*. Minneapolis, MS: IPUMS.

—, —, —, — and — (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.

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.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, **37** (2), 187–204.

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.

Rosenbaum, P. R. (2002). *Observational Studies*. New York: Springer.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J. and Sobek, M. (2020). *IPUMS USA: Version 10.0*. Minneapolis, MS: IPUMS.

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.

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.

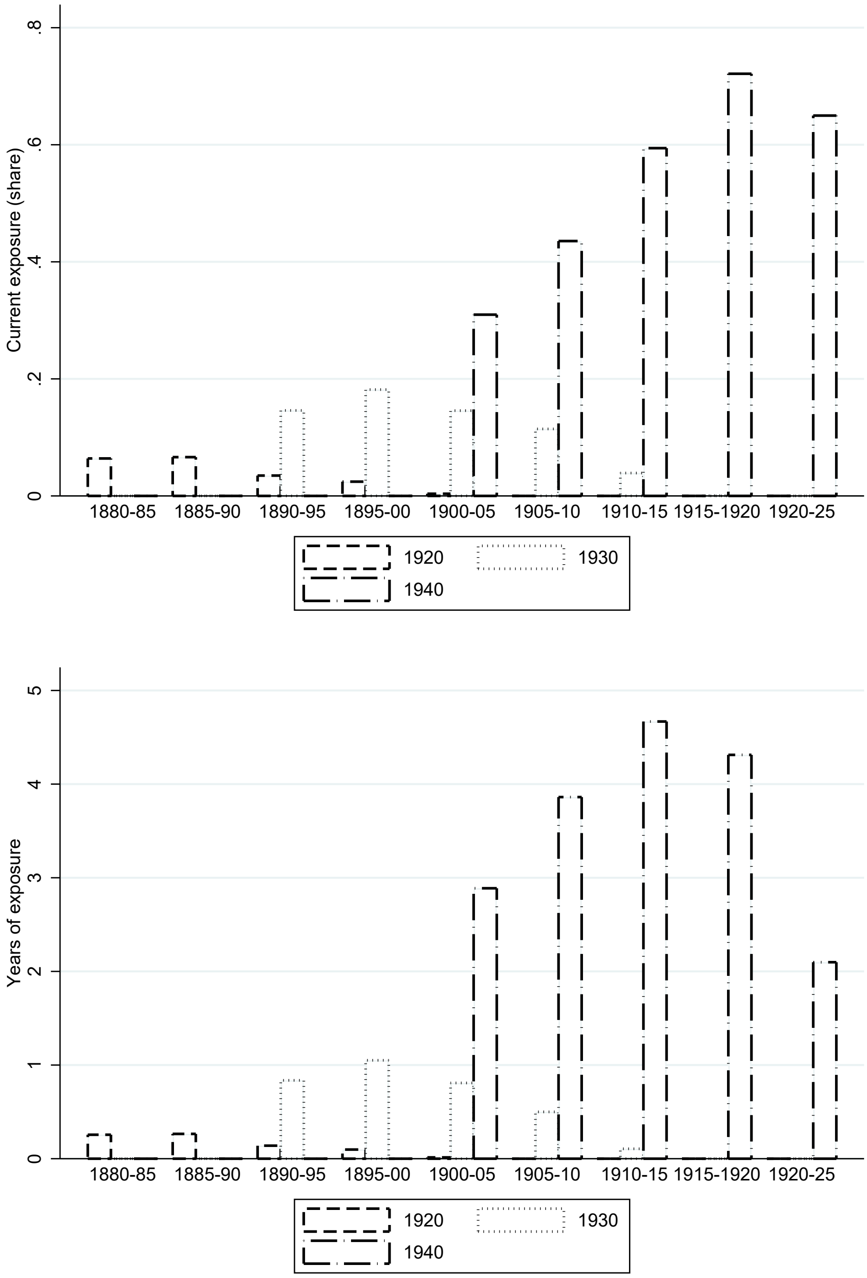
Wanamaker, M. H. (2012). Industrialization and fertility in the nineteenth century: Evidence from South Carolina. *The Journal of Economic History*, **72** (1), 168–196.

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

A. Web appendix

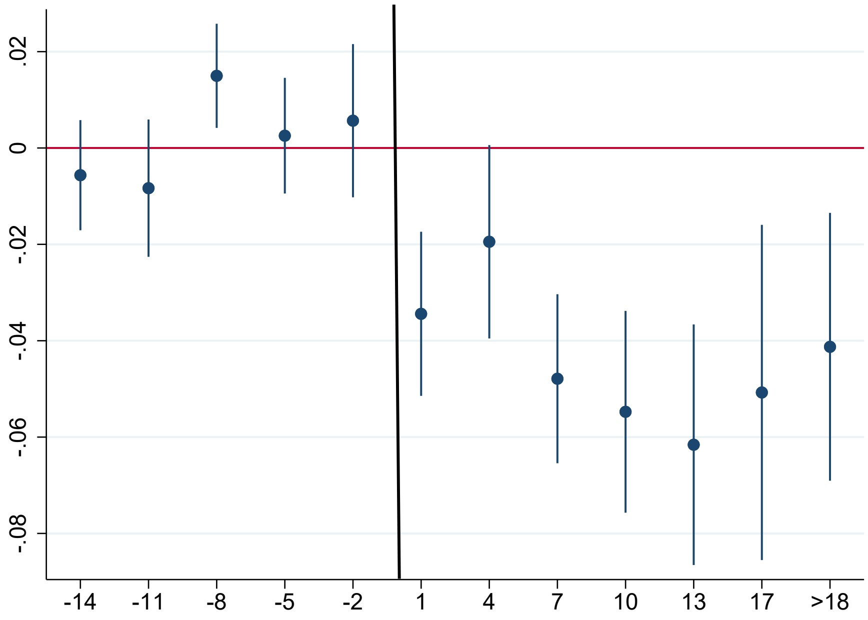
This Web Appendix (not for publication) provides additional material discussed in the unpublished manuscript " The Impact of Margaret Sanger's Birth Control Clinics on Early 20th Century US Fertility" by Stefan Bauernschuster, Michael Grimm, and Cathy M. Hajo.

Figure A.1 — Exposure to a birth control clinic by cohort and year



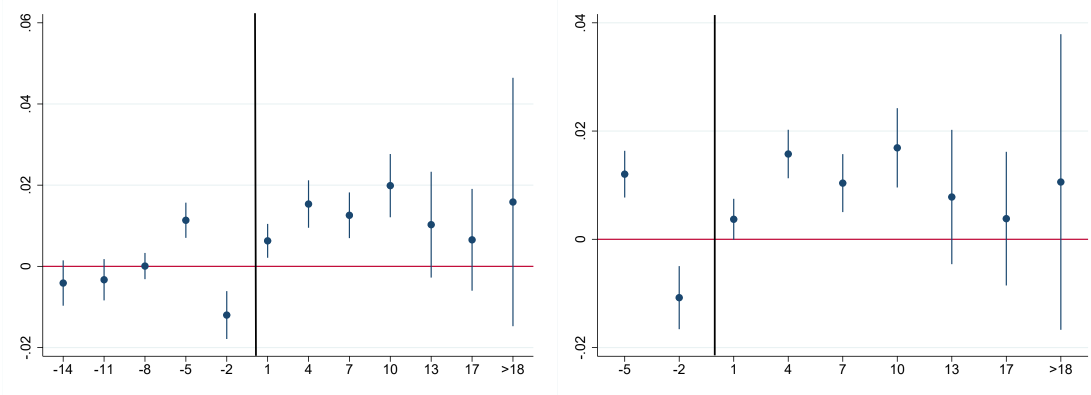
Notes: Data source: IPUMS US Census, 1920, 1930 and 1940.

Figure A.2 — Event-study plot of the impact of exposure to a birth control clinic on fertility, restricting the sample to women from ever-treated counties



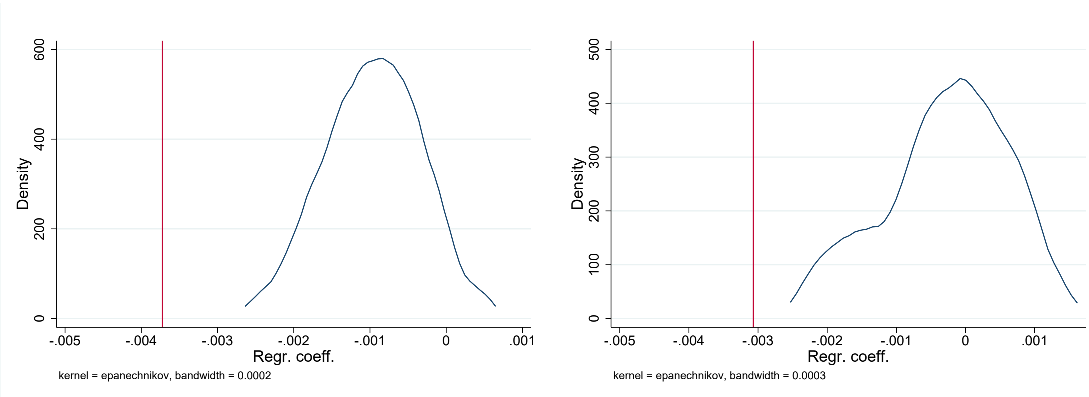
Notes: Data source: IPUMS US Census, 1920, 1930 and 1940. The figure shows event study estimates of the exposure to a birth control clinic on the number of children below the age of five in a household. Treatment effects are estimated along the lines of equation 2, but restricting the sample to women from ever-treated counties. The whiskers mark the 95 percent confidence band.

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



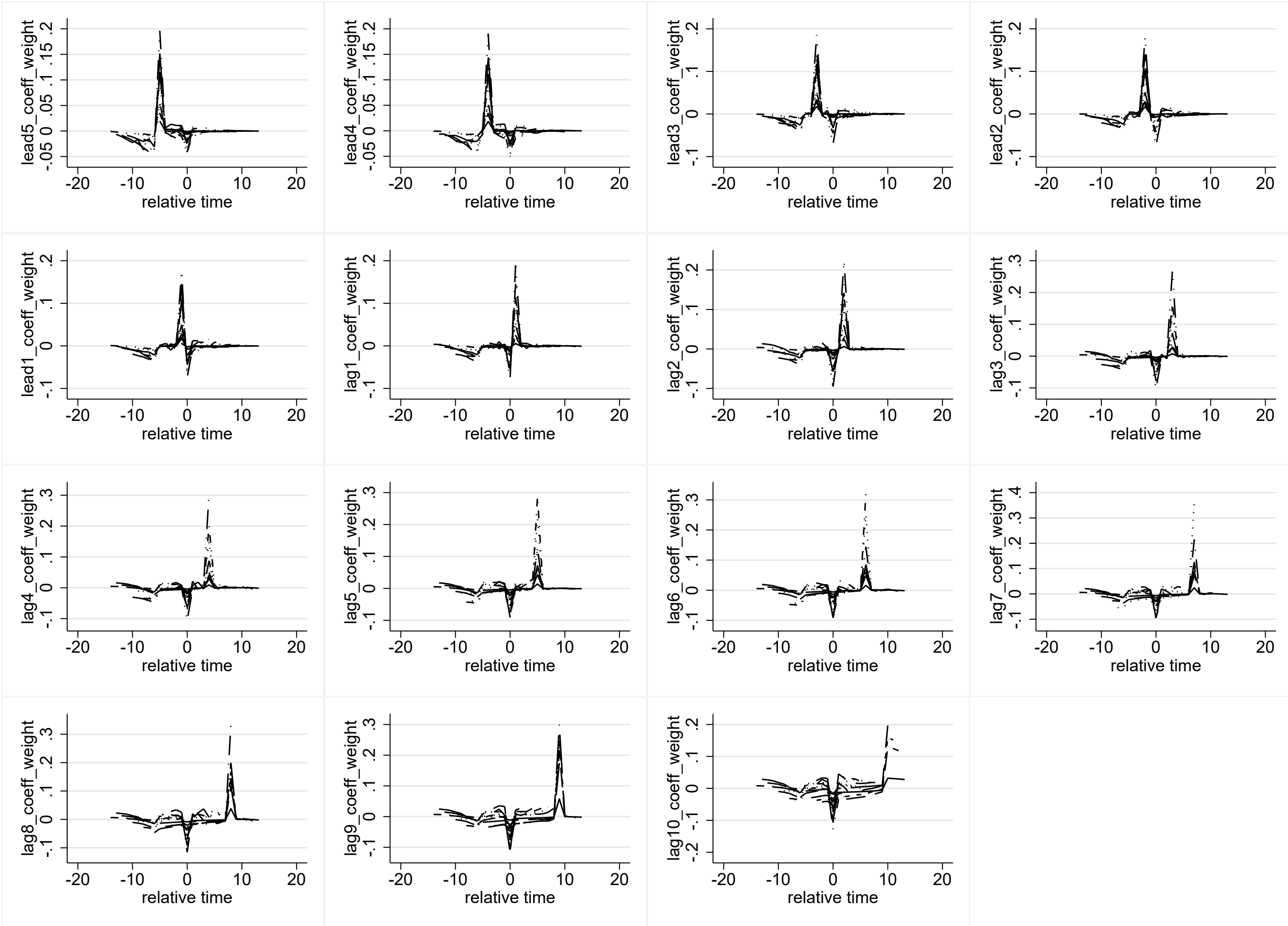
Notes: Data source: IPUMS US Census, 1920, 1930 and 1940. The figure shows event study estimates of the exposure to a birth control clinic on women's labor force participation (left panel) and women's employment status (right panel). Treatment effects are estimated along the lines of equation 2. The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

Figure A.4 — Randomization inference test



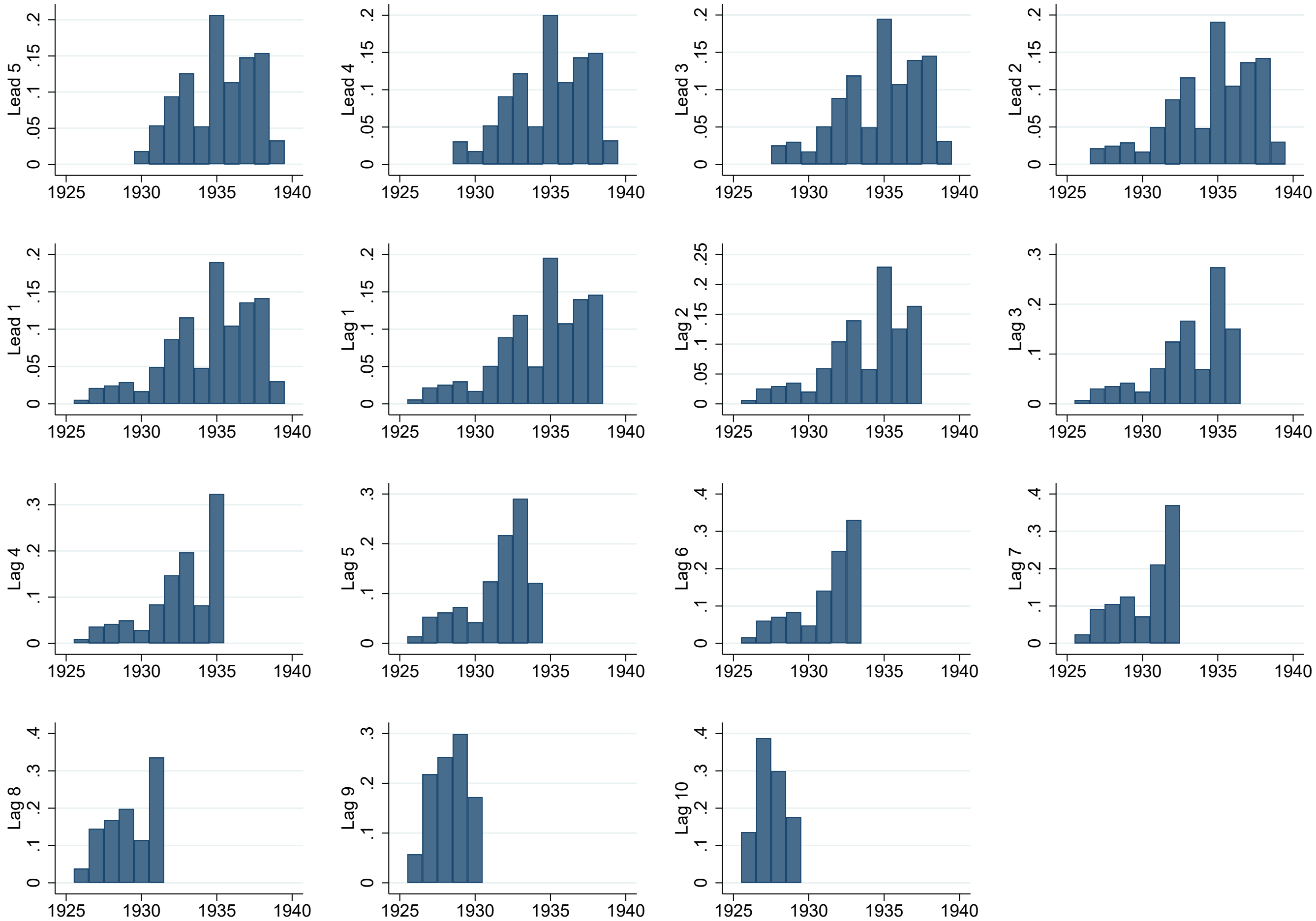
Notes: Data source: IPUMS US Census, 1920, 1930 and 1940. The figure shows randomization inference tests based on 100 permutations following Rosenbaum (2002) and Cohen and Dupas (2010) for the birth cohort by state fixed effects (left panel) and the county fixed effects model (right panel).

Figure A.5 — 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.

FIGURE A.6 — Event-study weights using Sun and Abraham (2021)'s 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 (2021)'s interaction weighted estimator.

TABLE A.1 — Socio-economic structure of counties with and without birth control clinics (BCC) in own or adjacent county over time

	1920		1930		1940	
	with BCC	w/t BCC	with BCC	w/t BCC	with BCC	w/t BCC
Age (mean)	30,049	29,051	29,845	29,083	29,151	28,940
Afro American (share)	0,016	0,111	0,065	0,120	0,103	0,108
Foreign born (share)	0,319	0,073	0,130	0,035	0,028	0,017
Literate (share)	0,949	0,941	0,976	0,967	0,971	0,965
Protestant (share)	0,314	0,650	0,504	0,655	0,648	0,650
Catholic (share)	0,493	0,201	0,359	0,196	0,215	0,192
Other Religion (share)	0,194	0,150	0,139	0,151	0,138	0,160
Farm HH (share)	0,115	0,484	0,168	0,461	0,352	0,464
Urban (share)	0,827	0,200	0,565	0,218	0,307	0,201
Big city (share)	0,632	0,016	0,197	0,015	0,047	0,002
Obs. (counties)	10	3022	121	2913	1358	1675

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. Birth control clinics statistics by Hajo (2010).

Table A.2 — The impact of exposure to a birth conrol clinic on fertility - detailed regression results

	(I)	(II)	(III)	(IV)
Years of exposure to BCC	-0.004***	-0.004***	-0.003***	-0.003***
	(0.001)	(0.000)	(0.000)	(0.000)
Age	0.165***	0.165***	0.166***	0.178***
	(0.002)	(0.002)	(0.002)	(0.002)
Age squared/100	-0.326***	-0.328***	-0.329***	-0.332***
	(0.003)	(0.003)	(0.003)	(0.003)
Residence in urban area (=1)	-0.232***	-0.139***	-0.111***	-0.111***
	(0.002)	(0.002)	(0.001)	(0.001)
County level fertility 1920			0.439***	
			(0.005)	
County level fertility 1920 x year 1920				0.582***
				(0.006)
County level fertility 1920 x year 1930				0.463***
				(0.005)
County level fertility 1920 x year 1940				0.310***
				(0.006)
African American (=1)		-0.148***	-0.128***	-0.130***
		(0.002)	(0.002)	(0.002)
Literate (=1)		-0.103***	-0.098***	-0.097***
		(0.005)	(0.005)	(0.005)
Foreign born (=1)		0.172***	0.168***	0.166***
		(0.005)	(0.004)	(0.004)
Protestant (county share)		Ref.	Ref.	Ref.
Catholic (county share)		0.115***	0.071***	0.072***
		(0.006)	(0.004)	(0.004)
Other religion (county share)		0.015	-0.008	-0.008
		(0.011)	(0.008)	(0.007)
Farm hh (=1)		0.176***	0.154***	0.153***
		(0.001)	(0.001)	(0.001)
Residence in big city (=1)		-0.058***	-0.015***	-0.016***
		(0.002)	(0.002)	(0.002)
State-specific cohort-fixed effects	yes	yes	yes	yes
R2	0.072	0.084	0.088	0.088
Observations	45,104,489	45,104,489	45,104,489	45,104,489

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. The table shows OLS regressions. The dependant variable is the woman's number of children below the age of five living in the household. Standard errors are clustered at the birth cohort by county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A.3 — The impact of exposure to a birth conrol clinic on female labor force participation and employment

	(I) Labor force	(II) Employ.
Years of exposure to BCC	0.001**	0.001**
	(0.000)	(0.000)
State-specific cohort FE	yes	yes
Age and urban area controls	yes	yes
Socio-economic controls	yes	yes
R^2	0.080	0.062
Observations	41,760,817	29,744,611

Notes: Data sources: IPUMS US Census, 1920, 1930 and 1940. The table shows OLS regressions. All regressions control for age and age squared as well as a variable indicating whether a houshold is in an urban area. 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 share of Catholics in a county. Standard errors are clustered at the birth cohort by county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4 — Matching treatment and control counties

	Probit	Unmatched sample		Matched sample	
	Treated	Treated	Control	Treated	Control
Birth rate 1925	0.324***	22.966	21.548	22.559	22.204
	(0.007)	(11.396)	(7.742)	(5.731)	(10.539)
Stillbirth rate 1925	0.005*	35.904	31.910	36.122	34.373
	(0.003)	(14.861)	(16.825)	(14.542)	(18.465)
Infant mortality 1925	0.004***	70.445	66.519	70.833	66.625
	(0.002)	(21.875)	(40.503)	(21.665)	(21.627)
Mortality rate 1925	-0.043***	10.068	8.596	10.099	9.222
	(0.013)	(3.272)	(4.544)	(3.224)	(5.887)
Foreign born (share 1920)	0.727	0.089	0.073	0.089	0.067
	(0.542)	(0.098)	(0.083)	(0.098)	(0.083)
Literate (share 1920)	-8.834***	0.954	0.965	0.954	0.955
	(1.204)	(0.046)	(0.053)	(0.046)	(0.062)
Farm households (share 1920)	-1.737***	0.391	0.518	0.389	0.483
	(0.239)	(0.228)	(0.194)	(0.225)	(0.208)
Blacks (share 1920)	-1.589***	0.074	0.060	0.073	0.089
	(0.476)	(0.128)	(0.147)	(0.127)	(0.178)
Age (average 1920)	0.512***	30.023	29.693	30.035	29.794
	(0.068)	(0.972)	(0.809)	(0.970)	(0.899)
Catholics (share 1920)	-0.844***	0.232	0.218	0.233	0.191
	(0.224)	(0.220)	(0.209)	(0.220)	(0.199)
Female labor (share 1920)	0.521	0.199	0.168	0.200	0.182
	(0.636)	(0.079)	(0.071)	(0.079)	(0.088)
Wage (average 1920)	-0.001***	1,006	1,019	1,005	993
	(0.000)	(256)	(275)	(256)	(263)
Urban county	-0.157	0.277	0.148	0.279	0.184
	(0.104)				
Big city county	1.326***	0.053	0.002	0.053	0.005
	(0.391)				
Observations 1925	1,757	905	916	890	385
Observations total		13,590	13,680	13,350	5,775
Pseudo $R2$	0.146				

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 (column 1) as well as a comparison of treated and control counties in the unmatched balanced panel 1925-1939 (columns 2 and 3) as well as a comparison of treated and control counties in the matched balanced panel 1925-1939.

Table A.5 — Descriptive statistics: Administrative data

	Balanced panel 1925-1939, matched	Balanced panel 1920-1939, matched	Balanced panel 1925-1938, no never-treated	Unbalanced panel 1920-1939
Birth rate	19.459	20.640	19.752	20.338
	(6.320)	(7.465)	(9.416)	(8.310)
Stillbirth rate	32.827	32.448	33.805	33.325
	(15.399)	(14.408)	(14.895)	(18.382)
Infant death rate	59.663	63.212	61.868	62.054
	(21.484)	(22.374)	(20.854)	(27.173)
All-age death rate	10.026	10.233	10.270	9.244
	(3.898)	(4.251)	(3.526)	(3.682)
Distinct counties	1,275	915	907	3,011
Obs.	19,125	18,300	11,791	48,515

Notes: Data sources: U.S. Vital Statistics. The table shows the means and standard deviations (in parentheses) of the variables in the sample that is balanced in calendar time from 1925-1939 (main sample), balanced in calendar time from 1920 to 1939, balanced in calendar time from 1925 to 1938 without never-treated, and on the full unbalanced panel from 1920 to 1939 available in U.S. Vital Statistics data base.