CESifo WORKING PAPERS

9627
2022

March 2022

The Effects of Incentivizing Early Prenatal Care on Infant Health

Kamila Cygan-Rehm, Krzysztof Karbownik

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9627*

The Effects of Incentivizing Early Prenatal Care on Infant Health

Abstract

We investigate the effects of incentivizing early prenatal care utilization on infant health by exploiting a reform that required expectant mothers to initiate prenatal care during the first ten weeks of gestation to obtain a one-time monetary transfer paid after childbirth. Applying a difference-in-differences design to individual-level data on the population of births and fetal deaths, we identify modest but statistically significant positive effects of the policy on neonatal health. We further provide suggestive evidence that improved maternal health-related knowledge and behaviors during pregnancy are plausible channels through which the reform might have affected fetal health.

JEL-Codes: I120, I180, J130.

Keywords: prenatal care, neonatal health, conditional cash transfers, prenatal care timing.

*Kamila Cygan-Rehm**
Leibniz Institute for Educational Trajectories
Wilhelmsplatz 3
Germany – 96047 Bamberg
kamila.cygan-rehm@lifbi.de

Krzysztof Karbownik
Department of Economics
Emory University
1602 Fishburne Drive
USA - 30322, Atlanta, Georgia
Krzysztof.Karbownik@emory.edu

*corresponding author

January 2022
Kamila Cygan-Rehm gratefully acknowledges funding by the Joachim Herz Stiftung. We thank Silke Anger, Pietro Biroli, Marianne Bitler, Reyn van Ewijk, Christina Felfe, David Figlio, Libertad González, Nabanita Datta Gupta, Hilary Hoynes, Regina T. Riphahn, Katharina C. Spieß, and seminar participants at Emory University, the University Erlangen-Nürnberg, the Ausschuss für Sozialpolitik in Halle, the IRLE Visitors Workshop at the UC Berkeley, the EEA 2020, the VfS Virtual Congresses 2020, the BeNA Workshop 2020, and the ESPE 2021 for helpful comments. We are extremely grateful to Andrzej Wojtyła and Cezary Wojtyła for providing access to the Pol-Prams data. Declarations of interest: none.

1. Introduction

Prenatal care is a central component of public policy since it targets both maternal and child health. Its main goals include improving maternal health knowledge, preventive monitoring of pregnancy, early diagnosis of high-risk pregnancies, maternal infections, and fetal abnormalities, as well as their potential treatment (see, e.g., Alexander and Korenbrot (1995) or Currie and Grogger (2002)). Specific regulations regarding the quality, content, and timing of medical care during pregnancy differ considerably across countries (Bernloehr et al., 2005); however, there seems to be a consensus that prenatal care should begin as early as possible (Alwan et al., 2016). For example, the World Health Organization (2016) recommends its initiation at the latest in the 12th week of gestation, arguing that the first trimester of pregnancy lays the foundations for a child's future development. However, despite the general increase in the coverage of early prenatal care over recent decades, it is still far from universal. Even in the highest-income countries, on average, only 82% of women initiated prenatal care in the first trimester of pregnancy in 2013 (Moller et al., 2017). In the US, in 2016, the overall rate of first-trimester care was 77%, with a substantial variation across states and socioeconomic groups (Osterman and Martin, 2018). On the other hand, visits during the first trimester have long been endorsed as means to prevent problems that may occur later in the pregnancy and to advise expectant mothers on beneficial behaviors, such as eliminating smoking, alcohol and illegal substance consumption (Bailey and Sokol, 2008), paying attention to healthy nutrition (Godfrey et al., 1996), taking the recommended vitamin supplements (Hovdenak and Haram, 2012), and using additional screening opportunities (Di Giacomo et al., 2022). However, to date, we do not know much about the causal link between the timing of prenatal care and child health and in particular about shifting its initiation to earlier in pregnancy. Little is also known about what programs and policy interventions might effectively incentivize early prenatal care utilization.

Most economists who study the link between prenatal care and birth outcomes draw on the theoretical framework by Grossman (1972), where prenatal care enters the "infant health production function" simultaneously with maternal health endowments, behaviors, and other inputs. Thus, a key empirical challenge is the potential endogeneity of the demand for prenatal care, which is likely correlated with unobserved maternal characteristics and other parental inputs that simultaneously affect infant health. Early empirical research addressed this endogeneity problem and estimated the production functions using a two-stage approach that used a set of individual-level and local-area instrumental variables, such as goods prices, health infrastructure, public expenditures, and labor market conditions (e.g., Rosenzweig and Schultz, 1983 or Grossman and Joyce, 1990). These studies generally find that prenatal care moderately improves neonatal health. More recent literature has increasingly used quasi-experimental designs that exploit sources of exogenous variation from changes in Medicaid policies (e.g., Joyce 1999; Gray, 2001; Currie and Grogger, 2002) or bus strikes (Evans and Lien, 2005). These studies focus primarily on socioeconomically disadvantaged mothers and the effects of expanded access to prenatal care and its utilization, rather than the timing of its commencement. Evans and Lien (2005) provide suggestive evidence that a loss of a prenatal visit early on in the pregnancy is more detrimental than visits lost later on, but their delineation happens only at the 6th month of pregnancy. Generally, the literature is silent on the possible gains from shifting prenatal care initiation to much earlier stages in the pregnancy.[1]

Recent literature reviews support these notions and conclude that we still know relatively little about the causal effect of prenatal care on child outcomes, and especially the role of its timing (Almond and Currie, 2011; Currie and Rossin-Slater, 2015; Corman et al., 2019). At the same time, prenatal care should be viewed as one of the earliest investments in a child's development, and growing evidence documents that health status at birth has long-term consequences for socioeconomic outcomes, such as education (Almond and Currie, 2011; Figlio et al., 2014), employment and wages (Black et al., 2007), mortality (Bharadwaj et al., 2017), and other adult outcomes (Almond et al., 2018). The question of whether the timing of prenatal care affects neonatal health becomes particularly important in light of self-productivity of human capital and dynamic complementarities of skills (Cunha and Heckman, 2007).

From a policy perspective, parallel to a discussion about the effects of prenatal care is the question of whether and how governments can encourage mothers to begin care as early as possible. While direct costs might create barriers in access to prenatal care in absence of a universal health care system, empirical evidence suggest that also in settings where pregnant women are eligible for free prenatal care, the utilization is still far from universal. For example, in the US, only two-thirds of women eligible for Medicaid had their first visit in the first trimester in 2016 (Osterman and Martin 2018). Two commonly suggested policy tools that could help to promote early initiation of prenatal care are information and monetary incentives. In the former case, governments or NGOs provide information to the public in the interest of changing a suboptimal behavior. For example, it has been shown that such interventions can be successful in educational (Wiswall and Zafar, 2015) and health (Keskin et al., 2017) decisions.

[1] Generally, there is little causal evidence on the effects of prenatal care from outside the U.S. Some evidence on the effectiveness of prenatal care access in lower income countries is provided in e.g., Gajate-Garrido (2013) for Philippines or Gonzalez and Kumar (2018) for Mexico. In the context of the U.S., Yan (2020) shows that within-sibling variation in prenatal care utilization is predictive of birth outcomes. However, none of these studies focus on the timing of prenatal care.

Providing only information or guidance, however, may not be sufficient since its reception and understanding may vary across different educational or socioeconomic groups (de Walque, 2010; Lange, 2011), or it simply may not induce the expected behavioral change (Prina and Royer, 2014).[2] Monetary incentives, in turn, have gained popularity in the developing world as an effective policy tool (Lagarde et al., 2007), and more recently, they have started being employed in developed countries (Leininger and Levy, 2015; Fryer, 2017).[3] Their main premise is a contract between a principal (government) and an agent (recipient) that payment will be made upon fulfilling a priori specified conditions.

In this paper, we evaluate the effects of a policy implemented in 2009 in Poland, which directly speaks to the questions of whether the timing of prenatal care matters and whether mothers respond to financial incentives. Specifically, the reform required mothers to obtain a medical certificate of having the first prenatal visit by the 10th week of gestation in order to be entitled to a one-time cash transfer paid after childbirth. Before the reform, no such certification was required, making the benefit unconditional. Given the universal coverage by public health insurance, the reform did not affect the eligibility for, the utilization of, or the access to prenatal care itself. Already before the reform, nearly all mothers used prenatal care services during pregnancy. Thus, for most mothers, the policy change effectively shifted the timing of their first prenatal care visit towards earlier in the pregnancy. Polish parliament passed the legislation on 6th December 2008 but the law was only effective for births on or after 1st November 2009. This unique hold-up in policy implementation allows us to separately identify the effects of providing unconditional public information versus a joint effect of monetary incentive and public information about the benefits of early initiation of prenatal care. We also focus on a setting where pregnant women are eligible for free prenatal care, so that there are no substantial direct monetary barriers in the access to health care services.

To estimate the effects of the reform on infant health, we use data on all births and fetal deaths in Poland between 2006 and 2011 paired with a difference-in-differences (DD) identification strategy. In that, we compare mothers whose conceptions happened shortly before and shortly after the policy-relevant cutoffs to mothers who conceived in the same weeks but in nonreform years. Overall, we observe statistically significant positive effects of the reform on neonatal outcomes, such as decreased fetal deaths and increased birth weight. On the other

[2] Buckles and Guldi (2017) suggest that information can be effective if it is provided to doctors rather than patients.
[3] Related more directly to birth outcomes, there is evidence from variation in taxation and benefits availability that expectant mothers respond to financial incentives when it comes to scheduling births (e.g., Dickert-Conlin and Chandra, 1999; Gans and Leigh, 2009; Tamm, 2012; Schulkind and Shapiro, 2014; Lalumia et al., 2015; Brunner and Kuhn, 2015; Borra et al., 2019).

hand, we do not identify economically meaningful effects on gestational age. We show that these results are not driven by potential selection issues and that the estimates are remarkably robust to alternative econometric specifications and sample restrictions. Due to the sequential nature of the legislation, we further document that financial conditionality is essential to generate the observed gains in neonatal health. Although relatively imprecise, our heterogeneity analysis points to somewhat larger effects among children from socioeconomically disadvantaged backgrounds and those at higher risks of adverse birth outcomes. Using auxiliary survey data, we also document post-reform improvements in maternal knowledge and behaviors related to drinking and smoking during pregnancy, which are consistent with the observed gains in neonatal health though not necessarily the exclusive mechanisms behind the positive effects. These behavioral changes are also larger for more disadvantaged mothers, which is in line with heterogenous effects of the reform on neonatal health.

These findings contribute to the existing literature in multiple ways. First, we show that the change in timing of prenatal care, even relatively early in pregnancy, affects infant health. This is different from prior work that focused on access to prenatal care and its utilization per se rather than the timing of its initiation. Thus, our results provide a validation of the prevalent medical recommendations (e.g., by the WHO) by exploiting a universal policy. Second, we document that health improvements from prenatal care can be achieved by introducing conditionality to an already existing unconditional family transfer granted after birth. Moreover, we show that public information stemming from the passage of the reform is itself insufficient to affect birth outcomes, and the conditionality of the cash transfer is a critical piece of the legislation that leads to improvements in neonatal health. To the best of our knowledge, this distinction between public information and conditional cash transfer has not been previously studied in the context of prenatal care and infant health. Documenting these phenomena is particularly important from a policy perspective given the inconclusive findings from other interventions previously studied in the literature (see, Corman et al., 2019). Finally, by exploring the potential channels behind the uncovered benefits of earlier prenatal care, we find improvements in maternal health-related knowledge and less hazardous behaviors during the pregnancy as plausible mechanisms.

2. Institutions

2.1 The Health care system and prenatal care in Poland

The Polish Constitution grants all citizens equal access to the publicly funded health care system, which is based on a mandatory insurance scheme.[4] Premiums are deducted from personal income, while those who do not work, irrespective of the reason, are fully subsidized by the state. Doctors working in the public health care system are state employees, and their wages are to a large degree centrally set. The majority of doctors work simultaneously in the public system and run their private practices, where patients pay out of pocket. In parallel, a limited number of private health care centers provide services almost exclusively to patients having an add-on private insurance. Due to universal coverage by the mandatory public health care insurance and an insufficient supply of doctors, in particular specialists, rationing of care in the public sector occurs through queuing and gatekeeping. A referral from a primary care doctor is, however, not needed to visit an OB/GYN specialist.

The excess demand in the public sector creates market space for private health care provisions. The private market has steadily grown from approximately 30 to 35 percent as a fraction of total health care expenditures between 2003 and 2010 (Łysoń, 2012). The private health care system comprises primarily services paid out-of-pocket, which are theoretically available to everyone. A much smaller segment of the market comprises private insurance add-ons, which refer to employer-sponsored insurance and are mostly provided by large firms. According to survey data from March 2009, nearly 10 percent of households had an add-on private insurance, but approximately 54 percent reported some out-of-pocket expenditures on health care services during the past three months (Diagnoza Społeczna, 2009).[5] Even though the private health care market has been growing during our study period, it was still uncommon to utilize it for more complex medical procedures (e.g., birth or cardiac surgery), and its primary focus was on consultations with specialists and dental treatments.

The standards of health care services are set and monitored by the Ministry of Health. In particular, prenatal care standards relevant for our study were issued in 2004 (Dziennik Ustaw, 2004) and stipulated eight to nine monthly routine consultations with a doctor or a midwife, depending on the duration of the pregnancy. Each of the visits included monitoring of maternal weight gain, blood pressure, general and gynecological examination, pregnancy risk

[4] Additional details are provided in Girouard and Imai (2000) or Nieszporska (2017).

[5] The average monthly spending in 2009 was of 175 PLN (approximately 55 USD at that time), which accounted for approximately 6.4 percent of the net household income. We calculated these numbers based on data from Diagnoza Społeczna (“Social Diagnosis”), which is a large biannual survey of a representative sample of the Polish population and has been recently used e.g., by Becker et al. (2020).

assessment, and promotion of a healthy lifestyle. In addition, these standards required several basic blood morphology tests, five fetal heart function assessments, three ultrasound screenings, and a test for gestational diabetes. The first visit was recommended to occur during the first ten weeks of gestation and it included some specific components such as a pap smear, a breast examination, VDRL/HIV tests, and in-depth pregnancy and lifestyle risk assessments including family history and social conditions. The in-depth risk assessment and promotion of a healthy lifestyle could be particularly sensitive to the first visit's timing.

Generally, in Poland, care during delivery and the postpartum period is of high quality, and almost all births (99.4 percent in 2008) are delivered in a hospital and attended by skilled healthcare personnel (UNICEF, 2021), out of which nearly 90 percent occur within the public healthcare system (NFZ, 2021) and the remaining 10 percent in private hospitals. In 2009, 27 percent of births in public hospitals were delivered via a caesarian section, and the median duration of a birth-related hospital stay was four days (NFZ, 2021). At the same time, almost all women used prenatal outpatient care services at least once during pregnancy (see Figure A1 in the Online Appendix).[6] According to various UNICEF indicators (UN IGME, 2021), Poland should be viewed as a developed, middle-income country with a high quality of health care, and thus our results could be generalizable to similar countries with publicly funded health care systems.

2.2 "Baby Bonus" and the Introduction of a Prenatal Care Requirement

In February 2006, the Polish government introduced a new one-time post-birth benefit, which aimed at increasing fertility and amounted to 1,000 PLN (ca. 318 USD or 262 EUR at the time). This was substantial as it constituted approximately 40% of average gross monthly earnings and 110% of the minimum monthly gross wage in full-time employment. This "baby bonus" was paid upon application and eligibility was universal, i.e., independent of family income or any other conditions.

A major change to the law was passed on 6th December 2008 with the explicit policy goal of promoting the early utilization of prenatal care. The new law stipulated that parents of newborns who applied for the "baby bonus" on or after 1st November 2009 needed to document that a mother's first prenatal visit was at the latest during the 10th week of gestation. The law

[6] We received the aggregate data from the National Health Fund (NFZ) of Poland (i.e., administrator of the public insurance scheme) based on a public information request. Unfortunately, the NFZ does not decompose the data by trimesters of pregnancy and the underlying individual-level data are not available for research purposes. In 2009, the coverage rate of 101% implies approximately a full coverage by prenatal care even after subtracting 11% for miscarriages because we need to add about 10% of privately-insured women (assuming that they use private health care providers).

defined a prenatal visit as a preventive consultation by a primary health care physician, a specialist in obstetrics/gynecology, or a hospital treatment during pregnancy.

The reform effectively introduced a financial penalty for not using early prenatal care and was motivated by an insufficient utilization of health services early in pregnancy, as well as by poor knowledge about healthy behaviors during pregnancy. For example, in the mid-2000s only 58.6 percent of women in urban and 46.0 percent in rural areas had their first prenatal visit during the first trimester of pregnancy (Kornas-Biela, 2012). This did not seem to improve over time as the numbers were at nearly the same levels already in 1997 (Druki Sejmowe, 1998). For comparison, since the 1990s, the trend was increasing in the U.S., and in 2003, almost 85% of births were to women who received first-trimester care (Corman et al., 2019). Most European countries had first-trimester rates of above 90% in the early 2000s (EURO-PERISTAT, 2008).[7] Nonetheless, in Poland, almost all expectant mothers used prenatal care at least once during pregnancy already before the reform, so that the policy did not aim at increasing the coverage rates but rather at accelerating the timing of the first visit (see Figure A1 in the Online Appendix).

The 10th-week requirement was passed unexpectedly as part of a law package that primarily focused on future maternity leave regulations.[8] To illustrate the public awareness, Figure A2 in the Online Appendix shows the development of the Google search index for "baby bonus" and "week of pregnancy" in Poland. The two vertical solid lines correspond to December 2008 and November 2009 – months when the reform was signed into law and implemented, respectively. The time series depicts no substantial increase in the search volume for "baby bonus" after the law passage in December 2008. However, the searches for "week of pregnancy" exhibit a sharp uptick right after the announcement of the new regulation.

Although the 10th-week threshold was linked to standards of prenatal care issued already in 2004, it became a controversial regulation from the beginning and was often criticized as too restrictive. One particular concern was related to the ability of the publicly financed health care system to cope with the expected shifts towards early prenatal visits and potential increases in total number of visits.[9] Thus, opponents argued that the reform could prolong waiting lines

[7] In 2004, the rates of first trimester prenatal care were as follows: Germany - 93.9%, France - 95.0%, Italy - 94.5%, Finland - 95.9%, and Sweden - 91.5%. Among countries of the former Soviet Bloc these numbers were: Czech Republic - 92.5%, Latvia - 91.8%, Lithuania - 74.5%, and the Slovak Republic - 79.5%.

[8] Specifically, the law passed on 6th December 2008 regulated an extension of maternity leave duration of 20 weeks by two additional weeks starting from 1st January 2010 and two further weeks starting from 1st January 2012. The law also enacted parental leave reserved for fathers of one week starting from 1st January 2010 and two weeks starting from 1st January 2012.

[9] Indeed, aggregate data do not exhibit any changes in the number of general practitioners or gynecologists and obstetricians under contract to the public system in 2008 or 2009 (see Figure A3 in the Online Appendix). At the same time, we find that the annual number of consultations and the daily consultation workload of OBGYNs

within the public system, which would prevent some mothers from obtaining the required certificate on time. Such capacity constraints in the timely access to the first doctor visit could have been particularly relevant for low-income mothers who were unable to afford any out-of-pocket medical expenses.[10] Otherwise, longer waiting periods in the public system would induce women to initiate prenatal care in private practice and to switch to a public health care facility as soon as possible.[11]

While the policymakers motivated the 10th-week threshold with late initiation of early prenatal care, there are only limited data on the timing of the first prenatal care visit in Poland. Unlike in the U.S. natality files, Polish birth records do not include any information on the timing of prenatal care initiation. Therefore, it is not possible to accurately estimate by how many weeks expectant mothers potentially sped up their first visit due to the reform. Nevertheless, the reform could not have affected the extensive margin of prenatal care utilization as virtually all mothers in Poland used prenatal health care services at least once during pregnancy already before the policy change. Still, to shed some light on the compliance with the new regulation, we draw on aggregate statistics on first-trimester initiation rates (i.e., up to 14 completed weeks of gestation). Figure 1 summarizes all available information from the publicly funded health care system. The data show that before the reform, on average, only 54% of expectant mothers initiated prenatal care during the first trimester of pregnancy and that this share was higher in urban than in rural areas. In contrast, in 2012, the average early initiation rate reached 79%, suggesting an increase of approximately 25 percentage points after the reform. It is important, however, to keep in mind that these numbers hide potential shifts

increased after 2008 as the new baby bonus criteria were introduced (see Figure A4 in the Online Appendix). Unfortunately, we do not have any evidence – either supported by the data or anecdotal – on whether this affected service quality. However, if the quality declined, our reduced-form estimates can be treated as a lower bound. Note that provision of health care including prenatal services is highly regulated in Poland, which to a certain degree mitigates the risk that overwhelmed physicians might undersupply the services. Furthermore, not all demand was concentrated among OB/GYNs as the policy allowed the first visit to be conducted by a general practitioner plausibly easing the demand on specialists.

[10] Approximate cost of a private visit during this time period was 80 to 100 PLN (36 to 45 USD at that time) and thus less than 10% of the "baby bonus". Therefore, from a financial standpoint, it even paid off to start prenatal care with a private provider to comply with the new legislation and then switch to free public health care system.

[11] After media reports on mothers who lost the "baby bonus" due to the reform, on 5th March 2010, parliament passed an amendment allowing for a medical certificate of at least one prenatal visit irrespective of its timing. Furthermore, mothers who lost eligibility between 1st November 2009 and 30th March 2010 could reapply for the "baby bonus". In January 2012, the tighter criterion of the first prenatal visit being at the latest in the 10th gestational week was reinstated. Finally, in 2013, a major reform abolished the universality of the "baby bonus" by introducing an income threshold for eligibility, which suppressed the share of eligible families to approximately 75% (see Figure A5 in the Online Appendix). Our focus in this paper is on the law change passed on 6th December 2008 and implemented on 1st November 2009, which was intended solely to promote early utilization of prenatal care. While it would be also interesting to study the policy amendment in 2010 and then its reinstatement in 2012, it is difficult to establish distinct treatment cutoffs for these later reforms and to find clear control groups (non-affected by the initial policy implementation).

towards earlier weeks within the first trimester itself. Moreover, the data do not cover visits with private health care providers, which account for approximately 20 percent of prenatal care visits in the first trimester.[12] Thus, the 25 percentage points increase is potentially the upper bound of the overall shift given that mothers using prenatal care within the public sector are of lower socioeconomic status, and thus potentially negatively selected in terms of early prenatal care utilization.

An indirect source of information on prenatal care utilization in the post-reform period in both the public and private sector is the take-up rate of the "baby bonus". Figure A5 in the Online Appendix presents governmental statistics on this issue. Not surprisingly, between 2006 and 2008, the take-up rates of the unconditional transfer were close to 100 percent. The coverage rate, however, remained remarkably stable in 2009 and 2010, when the eligibility was tied to a medical certificate of prenatal care. The decrease from 99.5 percent in 2008 to 99.1 percent in 2009 corresponds to approximately 1700 mothers who might have lost the benefit due to stricter legislation but in 2010, the coverage reached 99.9 percent again. This evidence strongly suggests that the vast majority of mothers complied with the 10^{th}-week threshold and the utilization of early prenatal care reached almost 100 percent during the post-reform period. A nearly full compliance is also evident in individual-level data from the Polish Pregnancy-related Assessment Monitoring System (Pol-PrAMS).[13] Between 2010 and 2012, approximately 98% of mothers declared that they had their first visit by the 10^{th} week of gestation, which confirms an almost universal compliance with the new regulation. According to the same data, on average, the first prenatal visit took place in the 6th week of gestation.

Taken together, we conclude that the reform did not affect the extensive margin of prenatal care utilization but it did affect the timing of the first visit.[14] This is not surprising

[12] There are no detailed data on private health care utilization, but the almost universal coverage rate of government payments after introduction of the conditionality in 2009 depicted in Figure A5 in the Online Appendix paired with approximately 80 percent early prenatal care utilization in the public sector depicted in Figure 1 leads us to believe that about 20 percent of women used private providers for their first prenatal visit after the reform. As we note in Section 2.1, at the same time, approximately 10% of households had an add-on private insurance, which fits the evidence that, in 2009, nearly 10% of women gave birth in private clinics. These women would also most likely use private providers for prenatal care from the beginning of the pregnancy. The remaining 10% needed to pay out of pocket for each private prenatal visit and most likely switched to the public system as soon as possible.
[13] We describe the Pol-Prams data in more detail in Section 8, where we discuss potential mechanisms behind the birth outcomes effects. Unfortunately, we are unable to reliably use the information about the timing of prenatal care initiation from the first survey in 2009.
[14] While the reform could not have affected the extensive margin of prenatal care utilization, shifting the first visit into earlier stages of the pregnancy potentially also leads to additional visits by the end of the pregnancy given the monthly schedule of recommended visits in the Polish prenatal care standards. Unfortunately, there is no information on the number of prenatal care visits in our main data to directly investigate this issue. Nevertheless, in Figure A4 in Online Appendix, we show a time series for the total number of public-sector consultations by specialists in obstetrics and gynecology to all women (not only those being pregnant). The figure illustrates an increase in visits after the reform, which we expect was driven by pregnant women. However, assuming the likely

given the cost of compliance was relatively low while the financial benefit quite substantial even if deferred in time.[15] Thus, our interpretation is that most mothers complied because they did not want to leave effectively "free money" on the table. On the other hand, the cash transfer itself did not have a scope for affecting birth outcomes directly, net of its effects on having a first prenatal care visit by the 10th week of gestation, because it was paid out after birth.

3. DATA

We use individual-level data from the Polish natality files for the years 2006-2011. The Central Statistics Office of Poland collects information on all of the approximately 400,000 live births and fetal deaths that occur in Poland each year from the original birth records. Given that fetal deaths are tracked after 22 weeks of gestation, the birth registry covers pregnancies that lasted at least 22 weeks.[16] The data files contain detailed demographic information on the mother, such as age at delivery, place of residence, number of previous births, marital status, education, and source of income. Children's characteristics comprise a distinction between live birth and fetal death,[17] birth date, birth weight, gestational length, and gender, but unlike, for example, U.S. birth records, there is no information on prenatal care utilization in the data. Importantly though, we do observe the exact day of birth and completed weeks of gestation, from which we obtain the date of conception.[18]

---

diminishing marginal returns over the number of visits, the additional visits late in the pregnancy should not be a major channel behind our reduced-form estimates.

[15] Corruption should not be an issue in our context since fully digitized system collecting information on all health care services billed within the public system was launched in January 2008. Given the high degree of transparency, monitoring, and control, it is unlikely that doctors backdated the first prenatal visit upon a women's request.

[16] The data also include a small number of pregnancies that lasted for 20 or 21 weeks. Nevertheless, we are unable to observe all pregnancies that terminated before the official threshold of 22 weeks either by a spontaneous or induced abortion. Early spontaneous abortion is associated with congenial defects rather than with socioeconomic factors (e.g. Currie and Grogger, 2002). In Poland, since 2000, their annual number was estimated at around 41,000, which translates to about 11% of pregnancies (see Figure A6 in the Online Appendix, panel B). Induced abortions are generally prohibited in Poland, except in three cases: if the pregnancy is a result of rape or incest, if the pregnancy threatens a woman's life, or if the fetus is irreparably damaged. Even then, an abortion needs to occur no later than in the 12th week of gestation, and a doctor might refuse to perform an abortion if it is not reconcilable with their conscience. Consequently, the number of registered abortions in Poland is negligible (see Figure A6, panel A). International organizations estimate the number of illegal abortions in Poland in a range of 1.5-5% of births (Zięba, 2006), but there are no reliable numbers on this issue. Nevertheless, various validity checks that we perform do not support the idea that selective abortion might be a threat to our empirical strategy. Furthermore, Figure A6 shows that official statistics on abortions and miscarriages do not exhibit any discontinuous patterns in years when the 10th week threshold was binding (i.e., 2009 and 2010).

[17] In the period under study, each pregnancy lasting at least 22 weeks, irrespective of the outcome, was registered using a standardized document, which was called "Birth Declaration" and is the source of our data. This document was filled in by professional healthcare personnel who was present during delivery or diagnosed fetal death. Unless medical stuff misclassifies "alive" vs. "dead" on the certificate, we should observe a universe of live births and fetal deaths for pregnancies over 22 weeks. We are not aware of any literature suggesting that the Polish data would be misreported (as this is the case for fetal deaths in the US). Even assuming some misreporting, to bias our results it would have to be correlated with the reform under study, which is highly unlikely.

[18] Rau et al. (2021) apply a similar approach to obtain conception week. We first compute the actual number of completed weeks spent in utero by subtracting two weeks from the gestation length since it is measured from the

Our outcomes of interest capture different aspects of neonatal health. Specifically, we considered a child's birth weight, which is a common measure of the underlying health of newborns. To ease the interpretation, we calculate a natural logarithm of birth weight (in grams). Following previous studies, we also include an indicator for low birth weight (less than 2500 grams) as a separate outcome since it is a leading risk factor for infant mortality, and those infants who survive are at elevated risk for many adverse outcomes (Corman et al., 2019). Babies with a low birth weight also tend to be born prematurely, which motivates us to also study gestational age (in weeks). Furthermore, given that early prenatal care might decrease the risk of fetal death (Currie and Grogger 2002), we likewise examine this margin of neonatal health. Unfortunately, the birth records do not contain any information about the specific cause of death. To address concerns of multiple hypothesis testing and to increase statistical power, we also estimate a summary index of neonatal health, which we construct using the first component of a principal components analysis (PCA) based on: an indicator for fetal death, birth weight in grams, and gestational age in weeks. We document the details on the PCA analysis in Table A1 in the Online Appendix.

We restrict our sample to singleton births (97.5%) as multiple births are born at lower birth weight for reasons unrelated to prenatal care (e.g., Figlio et al., 2014).[19] Furthermore, we focus on mothers between 15 and 50 years old at birth (99.99%) and omit a small group (0.14%) of foreign residents. Finally, given the reform's timing, we focus on children conceived at most up to four weeks before and up to four weeks after 6th December 2008 or 8th February 2009, and exactly the same weeks in the surrounding years 2005/6 – 2009/10.

Table 1 shows the summary statistics for our primary estimation samples. Further details of this table are explained in the Online Appendix. For comparison, column 1 shows the means for all live and stillbirths in Poland between 2006 and 2011. Importantly, these descriptive statistics are very comparable across columns, suggesting that a subset of mothers that we are using in our main analysis is broadly representative of the population of all births as well as confirming a balance of covariates across treatment and control groups.

first day of a woman's last menses, which typically occurs two weeks later. Then, we convert the period in utero into days multiplying the weeks by seven and subtract it from a child's birth date. By doing so, we assume that a delivery occurred exactly on the next day following the last day of the reported week of gestation, but our results are robust to assuming a mid-week delivery in the following gestation week (i.e., if we add three days to the period in utero measured in days). We present these results in Section 7.

[19] Our results are nearly identical when we expand the sample to include multiple births, and the policy itself did not address mothers of singletons and multiple births differently.

4. EMPIRICAL STRATEGY

We use the law change passed on 6th December 2008 in Poland as a source of exogenous variation in the timing of the first prenatal care visit and estimate its effects on neonatal health. Specifically, we exploit the fact that depending on the date of conception, the reforms quasi-randomly assigned women to the new requirement of the first prenatal visit during the first ten weeks of gestation in order to receive the financial benefit. Figure 2 summarizes the legislative timeline and treatment status based on the date of conception.

The hold-up in implementation of the reform provides us with two thresholds along the conception date that we examine using a difference-in-differences (DD) design. The first threshold, 6th of December 2008, refers to the passage of the law. The latter cutoff, 8th of February 2009, is determined by the day of the law's implementation (i.e., 1st November 2009), which applied to births, and the common assumption that typically a fetus spends 38 weeks in the uterus if we count from the day of conception.[20] Thus, for women who conceived a child on 8th February 2009, a physician typically calculated an expected due date on 1st November 2009, when the new law became binding. The first cutoff induced exogenous variation in public information about the importance of early prenatal care, while the second was related to real monetary losses in the case of non-compliance with the policy. Having established the two cutoffs, we use a child's date of conception inferred from the data to determine whether and how its mother was affected by the reform.[21]

Consequently, we consider three groups of mothers: those who conceived until 6th December 2008; those who conceived between 6th December 2008 and 7th February 2009; and finally, those who conceived on 8th February 2009 or later. The first group of mothers was not affected by any part of the reform.[22] The second group most likely knew that change in

[20] Note that common medical standards for establishing the expected birth date count the duration of a pregnancy from the first day of a woman's last menstrual cycle, which typically gives a 40-week gestational period. Nevertheless, a conception is possible only after ovulation, which usually occurs two weeks later, thereby giving the fetus typically only 38 weeks in the womb.

[21] As we describe in Section 3, we infer a child's date of conception from its exact day of birth and the actual gestational age at birth. Thus, due to a substantial individual-level variation in the actual length of the pregnancy, there is no mechanical one-to-one correspondence between the date of conception and the date of birth (i.e., a shift by 38 weeks for all children). We use the date of conception instead of birthdate to determine treatment status because the day of birth might be endogenous if, for example, the reform affected the duration of pregnancy.

[22] We assume that these women were untreated. However, to be more precise, mothers who conceived few weeks before 6th December 2008 were still below the 10th week of gestation when the reform was passed. Thus, they could still have changed their prenatal care utilization due to the information spawned by the new law if they reacted immediately. Alternatively, their prenatal care could have been negatively affected if there was a sudden excess demand due the reform and the preference was given to mothers who conceived after the reform (i.e., who would have experienced financial consequences from a non-compliance with higher likelihood). For these reasons, in a robustness test presented in Section 7, we backdate the first cutoff by 10 weeks to 27th September 2008 to assure that no women in the control group could have been affected by the new policy in the first ten weeks of

legislation was passed but they also anticipated that they would not be financially affected as their expected due dates were still before 1st November 2009. Nevertheless, these mothers might have positively reacted to the information about the benefits of early prenatal care stemming from the law passage spread by the media and the government. This creates a unique treatment group that likely experienced a public information shock but remained unaffected by the conditionality of the cash transfer. Since the actual length of pregnancy is uncertain (e.g., Jukic et al., 2013), mothers with calculated due dates approaching 1st November 2009 might have also changed their behavior regarding prenatal care to avoid the risk of losing the "baby bonus". Therefore, we expect that compliance with the new law among this group was increasing along the conception date, where mothers who conceived a few weeks after 6th December 2008 comply solely due to the information channel, while those who conceived a few weeks before 8th February 2009 also complied to avoid financial losses. Finally, the third group of women are those with expected due dates on or after 1st November 2009 who knew that they will need to provide an appropriate prenatal care certificate to obtain the "baby bonus" after birth.

In our empirical analysis, we estimate two sets of regressions intended to uncover either the joint effect of information and cash transfer or just the impact of public information. In the first case, we use both cutoffs and a donut-hole sample that compares health outcomes of children conceived shortly (i.e., up to four weeks) before 6th December 2008 and shortly (i.e., up to four weeks) after 8th February 2009 relative to children conceived in the same weeks in surrounding years. In the second instance, using a difference-in-discontinuities setting, we compare the outcomes of children conceived up to four weeks before and after 6th December 2008 relative to children conceived in the same weeks in surrounding years.[23] Following other studies with similar designs, we estimate our main results for different bandwidth choices between 1 and 4 weeks around the relevant cutoffs. At the same time, a priori, we prefer estimates using the widest bandwidth for two reasons. First, larger sample sizes mitigate potential concerns related to analyzing rare events such as the incidence of fetal deaths and low birth weight.[24] Second, wider bandwidth also reduce the potential impact of the measurement

gestation. Using this alternative control group does not lead to different conclusions. Thus, in the main analysis, we use the exact policy cutoff (as is common in the literature) rather than the lagged formulae.

[23] One could also consider exploring 8th February 2009 cutoff using differences-in-discontinuities setting, however, women who conceived a few weeks before this date can be considered as plausibly treated by both the information and conditionality of the benefit. This is due to potential financial losses had their birth occur later than expected and uncertainty in exact pregnancy length. For this reason, we view our donut-hole difference-in-differences approach as less problematic from interpretation standpoint given that it contrasts individuals who were not affected by the reform with those who were affected by both of its elements.

[24] For example, on average, we observe approximately only 30 fetal deaths per week out of more than 74,000 births, which implies that in case of narrow bandwidths, some random sample variability might lead to spurious effects of large magnitudes. Since both fetal death and low birth weight are rare events, likewise using linear

error in the exact date of conception. Irrespective of the exact setting, we estimate the following equation:[25]

$$Y_{it} = \alpha + \beta Reform_{it} + l_c + f_{dy} + d_{dw} + d_{dw} \times c + \gamma X_i + \varepsilon_{it} \tag{1}$$

where Y_{it} is an outcome of a child i conceived on a given day t. Our explanatory variable of interest is the dummy variable $Reform_{it}$, which is equal to one if a child was conceived on or after the particular cutoff day during the reform winter of 2008/9 (i.e., up to four weeks after 6th December 2008 or 8th February 2009). Our regressions further include the following fixed effects: cohort for each winter between 2005/06 and 2009/10 (l_c), day of year taking values from 1 to 365 for non-leap years (f_{dy}), and day of week taking seven values (d_{dw}). The cohort fixed effects (l_c) adjust for any cohort-specific effects being common to children conceived during a particular winter, such as the business cycle. The day of year (f_{dy}) and day of week (d_{dw}) fixed effects capture any unobserved effects associated with conceptions occurring during holidays, on weekends, or any specific day within a year more generally. For example, they account for the fact that our estimation samples might comprise holidays (e.g., Christmas or New Year's Eve) and that both our policy cutoffs were weekend days. The day-of-year fixed effects also flexibly isolate any seasonality effects that are common across cohorts. Therefore, we do not need a separate dummy variable to control for unobserved differences between children conceived in the weeks before versus after a particular cutoff. Our econometric specification, which closely follows Gans and Leigh (2009) and Borra et al. (2019), also includes an interaction term between the day of week dummies and cohort. We further include a set of control variables, X_i, accounting for various sociodemographic characteristics, such as maternal age at conception (linear and squared), dummies for her marital status, education, source of income, province of residence, rural areas, number of previous births, and child's gender.[26] Finally, ε_{it} is an error term. Following Gans and Leigh (2009) and Borra et al. (2019), we apply date-level clustered standard errors, which in our case, adjust for correlation among

probability models might not be appropriate. Therefore, we also replicated these analyses using logit, rare events logit, and complementary log-log models. The results remain qualitatively and quantitatively (based on marginal effects) very similar.

[25] A similar strategy has been used, e.g., by Lalive and Zweimüller (2009), Dustmann and Schönberg (2012), Cygan-Rehm (2016), and Huebener et al. (2019), to evaluate the impacts of parental leave reforms. Given that we use daily-level data, our model specification is particularly close to Borra et al. (2019), who examined the health consequences of speeding up births for non-medical reasons. Similar to their work, we also include a post-reform period (in our case only a single 2009/10 cohort) as a control group to difference our seasonality effects, which is not problematic so long as the seasonality pattern remains the same across cohorts. We test robustness of our results to this inclusion in Section 7.

[26] The reform we are studying coincided with the global financial crisis but this is unlikely to confound our treatment effects since we focus on births around a relatively narrow time bandwidth surrounding the policy cutoffs. Nevertheless, in Section 7, we show that our primary results remain unchanged if we additionally control for local unemployment rates, which serve as a proxy for the potential impact of the recession.

children conceived on a particular day.[27] In a robustness section, we also adjust our estimates for multiple hypotheses testing as well as present inference based on standard errors clustered at the month of conception level.

Given that virtually all mothers in Poland used prenatal care at least once during pregnancy already before the reform, the counterfactual situation does not include changes in the take-up of prenatal care per se. Instead, the policy change allows us to identify the effects of speeding up the first prenatal care visit either solely due to information signal to the public or due to the conditionality of the cash transfer. Indeed, the aggregate statistics discussed in Section 2.2 point towards higher rates of women initiating their prenatal care during the first trimester. Even without knowing the exact magnitude of the "first-stage" shift in the timing of the first visit, we can interpret our estimates as reduced-form effects of incentivizing prenatal care during the first ten weeks of gestation either through information signaling or financial incentives following the information component. In that, without the cash conditionality incentivizing women to start prenatal care by the 10th week of gestation at the latest, the counterfactual would be a later timing of prenatal care initiation.

The key identifying assumption is that net of any potential seasonal effects captured by the fixed effects outlined above, there are no other factors beyond the policy change that differentially affected children conceived before and after the policy-related cutoffs. This implies that the treatments by public information and conditional cash transfer were "as good as" randomly assigned by the reform and solely conditional on the conception date. We perform multiple validity checks to assess the plausibility of this assumption. We summarize these results below and discuss the technical details in the Online Appendix. Specifically, we document that children born up to four weeks before and after the policy-related cutoffs do not differ in their sociodemographic backgrounds, i.e., the covariates are balanced (see Table 1). We also do not find any statistically significant or sizable changes in the number of conceptions across the thresholds (see Table A2 in the Online Appendix), which suggests that there was no endogenous sorting of mothers into conceptions due to the reform. We find, however, a small increase in the number of births after the new law became effective, which contradicts endogenous accelerations of deliveries to avoid the new requirement and rather suggests that the reform slightly prolonged the average duration of pregnancies, which we view as a positive outcome. Moreover, we show that in the reform and the nonreform years, the number of

[27] To some extent, our empirical approach closely resembles a regression discontinuity design with a forcing variable effectively being a date of conception, which is another reason why we chose to cluster standard errors at day of conception level (Lee and Card, 2008).

conceptions followed similar seasonality patterns around the cutoffs (see Figure A7 in the Online Appendix). Finally, in Section 7, we present additional robustness checks and an event study analysis documenting that our results are not confounded by pre-existing trends. For this purpose, we extend the main sample to 15 weeks before and after the policy-related cutoff dates for each included cohort. We estimate a similar model specification as in Equation 1 but additionally interact the reform dummy with event-time indicators defined relative to the first cutoff date. Thus, akin to our main specification, in the event study, we use the non-reform cohorts as a control group. To increase precision, we aggregate the event times to three-week intervals. All these exercises support our quasi-experimental design.

5. Effects of the reform on neonatal health

5.1 Main results

We begin with estimating the total effect of the policy change, i.e., of public information and conditionality of the financial benefit jointly, by using a donut-hole difference-in-differences design. In that, we compare the health outcomes of children conceived up to four weeks before 6th December 2008 and up to four weeks after 8th February 2009 relative to children conceived in the same weeks in surrounding years. Each panel in Table 2 shows the results for one of the following outcomes: summary health index at birth (standardized), probability of fetal death, the log of birth weight, probability of low birth weight, and gestational age in weeks. The outcomes in panels B, C, and D are multiplied by 100 to enhance visibility. In each case, we present the point estimate on the reform indicator across different sample bandwidths from one to four weeks around the relevant cutoffs, as well as with and without individual-level controls.

We observe that the reform significantly increased the summary health index at birth by 2.0 to 3.5 percent of a standard deviation.[28] Although the effect sizes vary across the different bandwidths, the point estimates are statistically indistinguishable from each other as their 95 percent confidence intervals largely overlap. Nevertheless, we prefer the more conservative results from the wider bandwidths because the estimates in the narrower windows might be susceptible to empirical issues related to the rareness of some specific outcomes behind the health index (especially fetal deaths) or the measurement error in the conception date.[29] The

[28] This result is not sensitive to alternative specifications of the health index. Using z-scores as in Kling et al. (2007), we obtain estimates ranging from 2.1 to 3.6 percent of a standard deviation while using a weighted mean of the standardized outcomes as in Anderson (2008), we get effect sizes of between 2.2 and 3.7 percent.

[29] Our main results could also be affected by potential spillover effects if mothers who conceived right before the reform announcement have been somehow affected by its information component given that their pregnancies were still below 10 weeks of gestational age when the reform was passed. In Section 7, we address this issue by backdating the law's passage by 10 weeks. Irrespective of the exact bandwidth chosen, using this alternative control group definition consistently yields point estimates that are stable and very similar to those in the last

estimated effects on fetal deaths imply nontrivial reductions in the event probability of at least 25 percent relative to the pre-reform sample mean. We also identify significant gains in birth weight of at least 0.3 percent (or 9 grams) and a corresponding reduction in the probability of low birth weight by at least 6.8 percent relative to the pre-reform mean. In some specifications, we also find a significantly extended gestational age, which is consistent with the increased number of births after 1st November 2009 shown in Table A2 in the Online Appendix. Nevertheless, the effects on gestational age are relatively small.[30] We conclude that the reform led to statistically significant improvements in neonatal health, which were, however, not solely attributable to extended gestational age. All results are invariant to including a rich set of control variables, which supports a quasi-experimental design. Moreover, our conclusions hold if we adjust the inference for multiple hypothesis testing by using resampled p-values or the Romano and Wolf (2005a,b) adjustment. Our main conclusions also hold if we use standard errors clustered at month of conception level instead of clustering at the date level. We document the alternative inference results in Table A5 in the Online Appendix.

The hold-up in the implementation of the reform allows us to estimate the effect of the information channel separately from the conditionality of the cash transfer. In that, we explore the policy signing cutoff and compare individuals with and without information on the government's intentions. The results are documented in Table 3, which mimics the structure of Table 2. Here, we do not find robust evidence for positive effects of the public information about the benefits of early prenatal care stemming immediately from the law passage and the subsequent media coverage. In fact, some of the estimates are even wrong-signed but generally small and virtually never statistically significant. Only in the +/- 4 weeks sample, all coefficients are signed in the expected direction, some are statistically significant, and those tend to converge to the estimates reported in Table 2. This makes sense given how the reform was implemented and the fact that compliance with the new legislation must have been increasing along the conception date. In that, when we use wider bandwidths, we potentially have more mothers who started complying with the new requirement to avoid financial penalties (Figure 2). The information channel results in Table 3 are likewise robust if we adjust for multiple hypothesis testing or standard errors clustered at month of conception level (see Table A6 in the Online Appendix). Overall, we conclude that the direct benefits of the information from the

column of Table 2. This makes us more confident that the conservative results in column 4 are the most preferable but also reveals that elevated estimates at smaller bandwidths might be due to the aforementioned spillovers.

[30] We further investigated effects on log gestational age and an indicator for premature birth, but these results are qualitatively similar to what we present in Table 2 for weeks of gestation. At the same time, we observe reductions in post-term deliveries, but these estimates are only significant in the one- and two-week bandwidth samples.

law announcement and its subsequent media coverage are minuscule compared to having both information and conditionality of the transfer. This suggests that the signal about the importance of early prenatal care initiation sent to the public was not sufficiently strong before financial consequences kicked in.

5.2 Discussion of the magnitudes

The reduced-form coefficients presented above correspond to intention-to-treat (ITT) effects of incentivizing prenatal care during the first ten weeks of gestation. To assess the magnitude of our estimates, we first compare them to reduced-form impacts of other programs affecting neonatal health from related literature. For instance, Rossin-Slater (2013) finds that access to the Women, Infants, and Children (WIC) program in a neighborhood increased average birth weight by 0.8 percent, while Hoynes et al. (2015) show that an expansion in Earned Income Tax Credit (EITC) payments reduced the incidence of low birth weight by up to 5.2 percent.[31] Almond et al. (2011) further document that the introduction of Food Stamps led to about a 1% decrease in the fraction of births below 2,500 grams. Our estimates are also in line with Borra et al. (2019), who find that abolishing the Spanish "baby bonus" significantly increased the number of early inductions of births, which led to birth weight reductions of up to 15 grams. Thus, we view our estimates as not only plausible but also comparable in magnitude to what was found in previous studies on other determinants of neonatal health.

To compare our results with earlier research on prenatal care, which typically focused on the availability of care in the U.S., we rescaled our reduced-form effects by the approximate increase in first-trimester rates of 25 percentage points (see Section 2.2). Thus, scaling up our reduced-form estimates by factor 0.4 yields back-of-the-envelope estimates of the treatment on the treated (TOT) effects of increasing the utilization of early prenatal care by ten percentage points.[32] Consequently, based on the 4-week bandwidth estimates, our results imply that a 10-percentage-point increase in early prenatal care leads to birth weight gains of 0.14 percent and reductions in the probability of birth weight below 2,500 grams of 0.16 percentage points or 3.1 percent relative to the pre-reform mean. These estimates are on a smaller side compared to prior literature. However, earlier studies focus primarily on estimating the effects of changes on the extensive margin (access to prenatal care) rather than on shifts in the timing of the first

[31] Using the same reform, however, Dench and Joyce (2020) document that EITC might not be causally linked to neonatal health. On the other hand, in the Uruguayan context, Amarante et al. (2016) find 19 to 25 percent declines in incidence of low birth weight as a result of cash transfers to poor pregnant women. These effects are larger than our reduced-form coefficients.

[32] Given that the change of 25 percentage points refers to women using the public health care providers, the value is likely an upper bound on the first-stage effect. A more conservative estimate of e.g., 20 percentage points would result in a multiplier of 0.50 implying somewhat larger TOT effects.

visit, and it is reasonable to assume that the latter effect sizes could be smaller. For example, using Medicaid expansions that increased eligibility for free prenatal care, Currie and Gruber (1996) observe reductions in the incidence of low birth weight of up to 5.3 percent for a 10-percentage-point increase in Medicaid take-up due to extensions targeting low-income groups.[33] Gray (2001) demonstrates an even larger impact using the variation in Medicaid reimbursement rates for prenatal services paid to the physicians; his estimates imply that a 10-percentage-point increase in prenatal care reduces the relative risk of low birth weight by 6.6 percent.

While most previous studies have focused on birth weight effects for infants born alive, we also find significant effects on fetal deaths, which are often associated with maternal complications of pregnancy. If these conditions are detected early enough, the death of the fetus can be often prevented by proper treatment in utero or emergency delivery (Currie and Grogger 2002). Our results imply that a 10-percentage-point increase in early prenatal care utilization reduces the relative risk of fetal death by at least 10 percent for mothers conceiving around the policy cutoff. For comparison, Currie and Grogger (2002) observe substantially larger effects using various administrative measures that increase eligibility for Medicaid and reduce nonprice barriers to prenatal care utilization in the U.S. Their results suggest a 40-percent reduction in fetal deaths among black mothers if prenatal care increased by 10 percentage points. That Currie and Grogger (2002) find much larger mortality effects also makes sense given that their variation is much more in the access to prenatal care (especially for the disadvantaged populations) rather than in its timing, which is what our reform shifted.

Taking a longer-term perspective, using the estimates from Figlio et al. (2014) and Black et al. (2005), our results for birth weight would suggest at least 0.07 percent of a standard deviation improvement in cognitive development at school or a 0.14 percent increase in adult income. Compared to other human capital inputs, such as being assigned to a high value-added teacher (Chetty et al., 2014), we view these potential gains from early prenatal care as relatively modest.

Our results suggest that financial conditionality baked into the reform was essential to generating the observed gains in neonatal health. Before the reform, however, the state was already paying the unconditional benefits, and thus, this monetary dimension allows us to compute willingness to pay for prenatal health. Previous research on tax incentives and birth

[33] Overall, for targeted changes applied to low-income groups, they identify a 2.6 percent reduction in the incidence of low birth weight if eligibility for Medicaid increased by 10 percentage points. Since the take-up was not full, but at about 49 percent, this effectively implies a 5.3 percent decrease in the incidence of low birth weight.

timing suggests that parental private willingness to pay is in the range of 1.2 to 3.1 USD per gram of birth weight (Schulkind and Shapiro, 2014; Clarke et al., 2017). Our birth weight estimates suggest gains of at least 9 grams due to the reform, which should be further scaled by a factor of four since we expect that at most 1 in 4 mothers shifted their prenatal care earlier in pregnancy as explained above. The transfer had a value of 1,000 PLN (or 318 USD at the time), and thus our implied willingness to pay for birth weight is 8.8 USD per gram. This estimate is between previous estimates for private willingness to pay and public willingness to pay stemming from programs such as WIC or Food stamps (Clarke et al., 2017). Overall, we view our estimates from an incentivized prenatal health intervention in Poland as remarkably close to those in prior literature where either private or public monetary valuation of neonatal health was assessed.

From a fiscal perspective, the reform built upon an already existing birth-related benefit, and thus it did not require any additional transfers from the state. The administrative costs were limited to a simple verification of the medical certification during an application process, which remained unchanged otherwise. From mother's perspective, the additional payment was much larger than a cost of visit in a private sector, to the extent that due to potentially increased wait times some women had to switch to such care in other to avoid losing the "baby bonus", while very few mothers actually lost the benefit due to almost universal compliance (see Figure A5 in the Online Appendix). At the same time, one cost item that we are unable to quantify is the additional burden on health care providers stemming from potential increase in the demand for prenatal care. On the other hand, we might also expect that the resulting improvements in neonatal health will potentially lead to some long-run savings in health expenditures. Irrespective, we find the modest gains in neonatal health documented above as encouraging from a policy perspective given that the government simply added conditionality to an already existing family benefit.

6. HETEROGENEITY

Prior research suggests substantial heterogeneity in the potential gains from prenatal care (e.g., Currie and Grogger, 2002; Conway and Deb, 2005; Abrevaya and Dahl, 2008; Sonchak, 2015). Typically, the effects are concentrated among socioeconomically disadvantaged groups (e.g., African-Americans, unmarried teen mothers, or school dropouts), which are at a higher risk of both inadequate prenatal care and adverse birth outcomes. Although our policy in question was universal, for lower-income mothers, the "baby bonus" constituted a relatively larger boost to their incomes compared to higher-income mothers. Thus, we might expect a bigger change in the utilization of early prenatal care among socioeconomically disadvantaged groups, which should be reflected by disproportionately better birth outcomes.

We present heterogeneity results by maternal education, employment, and place of residence in Table A3 in the Online Appendix. We also stratify by maternal age and birth order, hypothesizing that first-time and teen mothers might benefit relatively more from early prenatal care due to its focus on promoting knowledge about the course of pregnancy and healthy prenatal behaviors. The analysis is based on our largest donut-hole estimation sample that yielded the most conservative estimates in Table 2. In each panel of Table A3, we extend the specification from Equation 1 with an interaction term between the reform indicator and a maternal characteristic of interest. The arguably most disadvantaged group represents a reference category and the interaction terms yields the excess losses or gains relative to this reference group. For comparison, panel A repeats our main results.

Panel B yields some evidence for an educational gradient in the reform’s effects on neonatal health. Specifically, the reduced probability of fetal deaths is primarily driven by mothers with less than high school education. We observe similar patterns for the remaining outcomes, but the differences are less pronounced and are usually statistically insignificant. The estimates by maternal employment (Panel C) and for rural and urban areas (Panel D) do not yield any notable differences across these groups, but we find some significant heterogeneity by maternal age at conception (Panel E). Specifically, teenage mothers drive the decrease in the probability of low birth weight. Finally, we do not observe any significant differences between first-born and higher-order children (Panel F). If anything, the point estimates suggest that higher-order births experience slightly improved birth outcomes over first-borns. Given that higher-order children receive relatively less prenatal care (e.g., Brenøe and Molitor, 2018), the reform could have incentivized particularly non-first-time mothers to use early prenatal care.

Earlier research by Abrevaya and Dahl (2008) suggests that the largest beneficial effects of prenatal care might be concentrated among fetuses most at risk for adverse birth outcomes. Therefore, we also consider the potential heterogeneity in the treatment effects along the distribution of (log) birth weight and our summary health index using unconditional quantile regressions, summarized in Figure A8 in the Online Appendix. We obtain strikingly large point estimates at the very bottom of the birth weight distribution. Beyond that, however, the point estimates are generally indistinguishable from the average effects and remain relatively constant and mostly statistically significant over the entire distribution.

Taken together, our heterogeneity analysis indicates slightly larger effects among children from socioeconomically disadvantaged backgrounds and those at high risk for adverse birth outcomes. However, the premiums for those more vulnerable children appear limited, and our estimates are generally underpowered to detect statistically significant differences across most groups and outcomes. Thus, the results should be treated with caution.

7. Robustness

So far, we documented that our main results are robust to including a rich set of background characteristics and are not confounded by changes in conceptions or births around relevant policy cutoffs. To further strengthen the causal interpretation of our estimates, Table 4 provides additional robustness checks. We focus on the +/- 4 weeks bandwidth since these are the most conservative estimates in Table 2; however, the results for the other bandwidths are likewise robust. For comparison, in panel A of Table 4, we repeat our main estimates.

Regarding potential confounding factors, we are not aware of any other contemporaneous policy changes that could have differentially affected neonatal or maternal health for conceptions up to four weeks before 6th December 2008 and up to four weeks after 8th February 2009. Nonetheless, the latter cutoff coincides with the introduction of the "baby bonus" three years earlier on 9th February 2006. Since the transfer was always paid postpartum it should not have affected birth outcomes of children included in our pre-policy control group. Nonetheless, to mitigate any remaining concerns about a potentially confounding impact of this earlier reform, in panel B, we exclude the 2005/2006 cohort.[34] The results remain very similar.

Following Borra et al. (2019), in the main analysis, we include both pre- and post-reform cohorts as a control group to difference out seasonality effects. This approach increases precision (due to a larger sample) and is not problematic so long as the seasonality patterns remain the same across cohorts (see Figure A7 in the Online Appendix). Nevertheless, in panel C, we exclude the post-reform cohort 2009/2010 to mitigate the concern that seasonality could have changed post-reform.[35] Again, our results remained substantively unchanged.

In the main specification, we use the policy announcement – 6th December 2008 – as a delineation determining the treatment through the public information channel because mothers who conceived on or after this date knew about the new regulation from the very beginning of the pregnancy. However, mothers who were below the 10th week of gestation on 6th December 2008 found out about the policy sometime during the first ten weeks of their pregnancy, and

[34] We also checked that introduction of the "baby bonus" did not lead to additional conceptions by conducting an analysis akin to what we present in Table A2 in the Online Appendix but for a 9th February 2006 cutoff. None of the estimates are statistically significant nor are they consistently signed.

[35] This exercise further alleviates a concern that our estimates are confounded by any anticipatory effect of the policy relaxation passed on 5th March 2010 as it is theoretically possible that women who conceived a child by the end of our sample period, which ends on 7th March 2010, could have already suspected that the 10th week requirement might be repealed soon.

theoretically, they could have changed their behavior due to this information. If so, our control group includes some potentially treated women. Thus, in panel D, we backdate the cutoff by 10 weeks (to 27th September 2008) so that no women in the control group could have initiated their prenatal care according to the new requirement in response to the policy announcement. Our conclusions are very similar to those from our preferred specification.

In Panel E, we apply a different assumption for computing the exact conception date. As explained in Section 3, for the main analysis, we assume that each pregnancy ends on the next day following the completed gestational age in weeks, which implies some measurement error in our treatment variable. Thus, in panel E, we alternatively add three days to each pregnancy duration, thereby assuming a mid-week delivery in the last (non-completed) gestational week. These estimates, except for the probability of fetal death, remain very similar.

We also investigate whether our estimates are sensitive to the inclusion of flexible time trends in conception date, which are meant to capture any trends in neonatal health over time, e.g., due to advances in the quality of health care or the accuracy of pregnancy tests. In panel F, we include quadratic trends, which yields very similar findings, and if anything, the estimates become more precise. While not shown in the table, we obtain similar results when we only include a linear trend in conception date or add higher-order polynomials (up to fourth-order).

As mentioned in Section 2, the 10th-week threshold for the first prenatal visit was passed in a bundle of laws that primarily regulated future maternity leave provisions, which became effective in January 2012. While these regulations happened beyond the sample period, our analysis covers some earlier changes in maternity leave. It is highly unlikely, however, that these earlier reforms confound our results because none of them coincided with the cutoffs for the 10th-week threshold along the conception date. We test this conjecture in Panel G by including indicators for the expected maternity leave duration at birth (between 16 and 22 weeks), which is based on a mother's expected due date. Our main results remain unchanged.

Another concern is that the 10th-week threshold was implemented during the Great Recession. Using Icelandic data, Olafsson (2016) shows that exposure to the crisis during the first trimester in utero had detrimental consequences for newborns' health. However, Poland weathered this period remarkably well and was the only economy in the European Union that avoided recession in its aftermath. Moreover, our empirical strategy should flexibly capture any post-crisis effects through the inclusion of cohort dummies. Nevertheless, to account for any differential effects of the shock across regions, in Panel H, we additionally control for the monthly unemployment rate at the province level. As expected, our results are not sensitive to the inclusion of this additional variable. To validate our estimates, in panel I, we also evaluate the effects of a placebo reform within the same empirical design. In that, we move back the two reform-related cutoffs by six months to 6th June 2008 and 8th August 2008. Reassuringly, these placebo estimates are all close to zero and statistically insignificant.

Finally, in Table A4 in the Online Appendix, we present results from an event study analysis, which allows estimation of the effects beyond the narrow windows around the reform cutoffs and testing of the internal validity by investigating the pre-trends. The results do not reveal any sizable or statistically significant coefficients in the pre-period, with one exception in column 4; however, this estimate strongly departs from its surrounding coefficients, indicating it is likely a statistical aberration. In contrast, in the post-period, and especially in the weeks after the reform's implementation, we observe mostly statistically significant estimates that confirm our findings from Table 2. A notable exception to this pattern represents the effect on fetal mortality, which is concentrated in the first three weeks after the reform's implementation. Nevertheless, even if temporal, the reduction in infant deaths should be viewed as a positive finding. The remaining effects seem to persist beyond the narrow window around the policy implementation.

We conclude that our results are not sensitive to excluding particular cohorts from the control group, correcting for the potential measurement error in the assignment of the treatment, or changing the preferred model specification. Likewise, the data do not exhibit any worrisome pre-trends, and we do not find any effects using placebo cutoff dates. This supports the causal interpretation of our main results.

8. POTENTIAL MECHANISMS

To shed light on some potential channels through which the reform might have improved neonatal health, we rely on auxiliary data from the Polish Pregnancy-related Assessment Monitoring System (Pol-PrAMS). This survey is conducted during one defined week of the year in all Polish hospitals among new mothers in the first few days postpartum (typically 1-4 days after birth).[36] The data include information on health-related knowledge and behaviors during pregnancy but initially the focus was on smoking and alcohol consumption (Wojtyła and Wojtyła-Buciora, 2016). The first survey took place in June 2009, i.e., more than six months after the reform was passed. Importantly, however, the vast majority of interviewed mothers conceived their newborn in September and the first half of October 2008, i.e., about two months before the law's announcement. Thus, we view their behavior early in pregnancy as unaffected

[36] Because some hospitals and some mothers refused to participate in the study, the final sample covers approximately 80% of women who gave birth in Poland on the survey day each year (ca. 3,000 mothers).

by the reform.[37] In 2010, the data were collected in the second week of August, which implies that the second survey covers mothers who conceived in November 2009, i.e., shortly after the 10th-week threshold became a binding requirement for "baby bonus" eligibility. We also include two later surveys from November 2011 and March 2012.

These data allow us to look at changes in maternal health-related knowledge and behaviors using a before-after design comparing mothers surveyed in 2009 (control group) to those surveyed in the years 2010-2012 (treatment group). We estimate the following equation:

$$Y_i = \alpha + \sum_{k=2010}^{2012} \beta_k Post_i^k + \gamma X_i + \varepsilon_i \qquad (2)$$

where Y_i is an outcome of a woman i, while the post-reform changes of interest are captured by a series of three dummy variables $Post_i^k$ which are each equal to one if a woman was interviewed in the 2010 to 2012 surveys, respectively. If the reform altered maternal behaviors and knowledge, then we would expect statistically significant coefficients on β_k. Furthermore, if the estimates are indeed driven by the reform rather than just secular trends in these outcomes, we would expect all three coefficients on the $Post$ dummies to be similar in magnitude.[38]

We do three things to ameliorate, at least to some extent, concerns that the results are driven by other changes over time. First, we estimate both the unconditional changes, as well as a specification that controls for a rich set of maternal characteristics (included in X). Similar to our main analyses, we control for maternal age, marital status, education, province of residence, town size, number of previous births, and child's gender. Second, we further control for the number of health care providers and medical practices at the province-by-year level to net out any potential changes in health care supply over time. Finally, we use a placebo outcome – exposure to smoking at work – that arguably should not have been directly affected by the policy but has been shown to affect birth outcomes (Bharadwaj et al., 2014; McGeary et al., 2020). Thus, if we find any significant changes in exposure to smoking at work over time, they should reflect general trends in healthy behaviors, which would imply that our before-after comparisons for other outcomes are likely confounded.[39]

---

[37] These mothers could not have reacted by altering their early prenatal care in response to the reform; however, it is possible that they reported better knowledge and behaviors after birth because they were affected by changing social expectations due to public discussion related to the reform. Thus, if anything, these estimates should be treated as the lower bound of changes in behavior and knowledge.

[38] From an identification standpoint, to interpret β_k as causal effects, we needed the outcomes to be the same in 2009 and in later years absent the reform requiring the 10th week prenatal care certification. In that, no other factors affecting the outcomes could have changed across these years.

[39] During the years we consider about 30% of Polish adults smoked (with approximately 85% of these regularly). Poland introduced a general smoking ban in public places, which mainly applied to pubs, restaurants and public outdoor facilities, on November 15th, 2010. Since the law was passed on April 8, 2010 it should not affect either our main analysis based on birth records or the 2010 estimate in Table 5. The change in sign between 2010 and 2011-2012 estimates could be partially affected by the regulation but these estimates are small and statistically insignificant nonetheless. Smoking ban in most indoor workplaces (e.g., offices and factories) was in place since

Table 5 shows these results. The outcomes in columns (1) to (5) are associated with health-related knowledge, while those in columns (6) to (8) with risky behaviors during pregnancy. Column (9) investigates passive smoking at the workplace, which we view as a placebo outcome. All outcomes are indicator variables. Panel A shows the means of the outcomes for mothers who conceived in the pre-reform period. For example, before the reform, approximately 88 percent of mothers knew that small alcohol amounts during pregnancy are harmful to maternal and fetal health. At the same time, nearly 13 percent of them admitted to alcohol consumption during pregnancy.

Panel B shows the unconditional before-after differences compared to the baseline mean. The estimates remain remarkably stable when we add a rich set of demographic controls and control for health care supply in Panel C. Generally, we find statistically significant improvements in the awareness that alcohol consumption and smoking are harmful to maternal and a newborn's health (columns 1 to 5). This appears to be accompanied by reductions in admitted drinking and smoking during pregnancy (columns 6 to 8). Such changes are consistent with increased exposure to early prenatal care, which usually focuses on educating women about the course of pregnancy and promoting healthy behaviors.[40] The estimates are very similar across all post-reform years, which mitigates the concern that we are picking up secular improvements in health knowledge and behaviors.[41] Furthermore, in column 9, we do not observe any sizable or statistically significant changes in exposure to passive smoking at work, which likewise suggests that our estimates do not simply reflect trends in healthy behaviors.

Assuming that the estimates in Table 5 indeed reflect changes induced by the 10th-week policy, improved maternal behavior during pregnancy seems to be a plausible channel through which the reform could have affected neonatal health. This would be in line with prior work on the detrimental effects of smoking and drinking in pregnancy on birth outcomes (Abrevaya,

1995 but anecdotally the locations designated for smoking in such settings were not always properly separated. Together these factors could explain why as much as 11% of women in 2009 wave of Pol-Prams declare that they were exposed to smoking at a workplace.

[40] Here we refer to the applicable Polish standards of prenatal care services (Dziennik Ustaw, 2004).

[41] In principle, we could expect somehow smaller coefficients in 2011 and 2012 if women responded to the subsequent amendment and abolition of the 10th week threshold in a reversed way as to its introduction. However, it is not clear whether such "symmetric" response actually took place given that the initial policy likely changed the social norms regarding the timing of prenatal care, especially that the re-introduction of the 10th week requirement in 2012 was already announced in March 2010. This intuition is supported by the fact that the early prenatal care initiation rates remained remarkably stable after 2012 (Figure 1), despite the fact that about 25% of mothers did not have to comply with the 10th week threshold anymore because they lost the eligibility for the "baby bonus" when it became a means-test benefit (see Figure A5 in the Online Appendix). The Pol-PrAMS data also confirm that between 2010 and 2012 about 98% of mothers consistently indicated that they started prenatal care before the end of the 10th week.

2006; Fertig and Watson, 2009; Bharadwaj et al., 2014).[42] Specifically, our results in column 7 imply that the rate of maternal smoking during pregnancy dropped by, at most, 7.1 percentage points. Extrapolating this estimate through the lens of Abrevaya (2006) would imply increases in birth weight in the range of 9 to 13 grams, which is on the lower end of our effects in Table 2. Our point estimate on a decrease in drinking during pregnancy is at most 5.4 percentage points (column 6 in Table 5), which when paired with small effect sizes reported by Fertig and Watson (2009), can explain relatively little of the potential effect of a decrease in drinking on observed improvements in birth outcomes. Nonetheless, we view both of these channels as plausibly jointly contributing to the observed gains in health at birth.

Overall, the magnitudes implied by the above extrapolation exercise seem reasonable as it is inconceivable that all benefits of early prenatal care for birth outcomes accrue solely through reductions in drinking and smoking. Although we are reluctant to draw strong causal conclusions based on the estimates in Table 5, they are remarkably consistent, irrespective of the exact outcome or specification. Furthermore, in Table A7 in the Online Appendix, we show a heterogeneity analysis along similar dimensions as the one presented in Table A3 but for the health-related behaviors and knowledge. The patterns point towards stronger responses among the more disadvantaged groups, i.e., teenage mothers, lower educated mothers, and mothers in rural areas.[43] These results are consistent with our subgroup analysis for the effects on neonatal health, which we discuss in Section 6.

Generally, the complementary results from the Pol-Prams data provide suggestive evidence that prenatal care early in the pregnancy could promote healthy behaviors and represent a plausible, but not necessarily exclusive, channel through which the reform improved newborns' health. Unfortunately, we are not able to investigate the role of other potentially beneficial aspects of early prenatal care such as, e.g., advice on proper nutrition, recommended vitamin supplements, as well as early diagnosis of high-risk pregnancies, maternal infections, and fetal defects, and their timely treatment (when possible and necessary).

9. Conclusions

This paper investigates whether financial incentives for early initiation of prenatal care improve neonatal health. Specifically, we evaluate a policy intended to speed up the first prenatal care

[42] Nilsson (2017) provides compelling evidence from Sweden that prenatal alcohol consumption negatively affects the fetus thought adulthood and leads to worse labor market and educational outcomes. Although we cannot observe long-term outcomes in our data, this evidence suggests that our policy in question may have positive effects on children's success in adulthood.

[43] Unfortunately, there is no information on maternal employment status in the first Pol-Prams survey so that we cannot execute this specific split.

visit in Poland, which has a universal, publicly funded health care system. The reform forced expectant mothers to begin prenatal care during the first ten weeks of gestation to be eligible for a one-time monetary transfer after birth, which was an unconditional benefit before. Administrative and survey data on prenatal care utilization as well as on governmental payments suggest that the vast majority of mothers complied with the new requirement of early initiation of prenatal care.

Applying a differences-in-differences design to individual-level register data on the population of births and fetal deaths, we find statistically significant positive effects of the policy on neonatal health. These benefits are generated by mothers who were treated both by the conditionality of the monetary transfer and the public information about the importance of early prenatal care spread by the media and the government due to the law passage. In contrast, by exploring the hold-up in policy implementation, we do not identify any statistically significant or economically meaningful effects from the public information channel alone. Our estimates imply moderate gains in birth weight and significant but temporal decreases in the likelihood of fetal death. Given the mixed evidence on the benefits of prenatal care (Corman et al., 2019), our results are promising as the reform shifted only the timing of the first prenatal visit rather than affecting the extensive margin of healthcare utilization. Although imprecise, our heterogeneity analysis points to slightly larger benefits among children from socioeconomically disadvantaged backgrounds and those at higher risks of adverse birth outcomes. Using auxiliary survey data, we show that improved maternal health knowledge and less risky behaviors during pregnancy might be plausible, though not necessarily exclusive, mechanisms underlying the positive effects of early prenatal care initiation on neonatal health.

In summary, our results suggest that children in utero may benefit from earlier initiation of prenatal care. Given the existing evidence that children from socioeconomically disadvantaged backgrounds are at a higher risk of both inadequate prenatal care and adverse birth outcomes, our results also imply that the Polish policy could have reduced early-childhood inequalities. Moreover, it is important to keep in mind that public policies that aim at promoting prenatal care might have additional impacts far beyond neonatal health. For example, Miller and Wherry (2019) demonstrate that children who gained Medicaid coverage in the womb and during infancy exhibited better health and educational outcomes as adults. Related research also found that first-trimester care positively affects maternal health-related behaviors measured postpartum (e.g., Conway and Kutinova, 2006; Reichman et al., 2010) and that benefits of improved prenatal health persist even beyond the treated generation (East et al., 2019). From the policy perspective, it is also important to highlight that the Polish legislation built upon an already existing birth-related benefit, which seems less costly compared to developing a new family transfer aimed solely at increasing prenatal care. Therefore, in this case, even modest gains in neonatal health should be viewed as encouraging.

REFERENCES

Abrevaya, J. (2006). Estimating the effect of smoking on birth outcomes using a matched panel data approach. *Journal of Applied Econometrics*, 21(4), 489-519.

Abrevaya, J., Dahl, C. (2008). The effects of birth inputs on birthweight: Evidence from quantile estimation on panel data. *Journal of Business & Economic Statistics*, 26(4), 379-397.

Alexander, G. R., & Korenbrot, C. C. (1995). The role of prenatal care in preventing low birth weight. *The Future of Children*, 103-120.

Alwan, N. A., Roderick, P. J., Macklon, N. S. (2016). Is timing of the first antenatal visit associated with adverse birth outcomes? Analysis from a population-based birth cohort. *The Lancet*, *388*, S18.

Almond, D., Currie, J. (2011). Human capital development before age five. *Handbook of Labor Economics*, Vol 4B.

Almond, D., Hoynes, H., Whitmore Schanzenbach, D. (2011). Inside the war on poverty: The impact of food stamps on birth outcomes. *Review of Economics and Statistics*, 93(2), 387-403.

Almond, D., Currie, J., Duque, V. (2018). Childhood circumstances and adult outcomes: Act III. *Journal of Economic Literature*, 56(4), 1360-1446.

Amarante, V., Manacorda, M., Miguel, E., Vigorito, A. (2016). Do cash transfers improve birth outcomes? Evidence from matched vital statistics, and program and social security data. *American Economic Journal: Economic Policy*, 8(2), 1-43.

Anderson, M. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of thee Abecedarian, Perry Preschool, and Early Training Project. *Journal of the American Statistical Association*, 103(484), 1481-1495.

Bailey, B., Sokol, R. (2008). Pregnancy and alcohol use: Evidence and recommendations for prenatal care. *Clinical Obstetrics and Gynecology*, 51(2), 436-444.

Barreca, A., Page, M. (2015). A pint for a pound? Minimum drinking age laws and birth outcomes. *Health Economics*, 24(4), 400-418.

Becker, S. O., Grosfeld, I., Grosjean, P., Voigtländer, N., Zhuravskaya, E. (2020). Forced migration and human capital: evidence from post-WWII population transfers. *American Economic Review*, 110(5), 1430-63.

Bernloehr, A., Smith, P., Vydelingum, V. (2005). Antenatal care in the European Union: a survey on guidelines in all 25 member states of the Community. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *122*(1), 22-32.

Bharadwaj, P., Johnsen, J., Løken, K. (2014). Smoking bans, maternal smoking and birth outcomes. *Journal of Public Economics*, 115, 72-93.

Bharadwaj, P., Lundborg, P., Rooth, D. O. (2017). Birth weight in the long run. *Journal of Human Resources*, 53(1),189-231.

Black, S., Devereux, P., Salvanes, K. (2007). From the cradle to the job market? The effect of birth weight on adult outcomes of children. *Quarterly Journal of Economics*, 12 (1), 409-439.

Borra, C., González, L., & Sevilla, A. (2019). The Impact of Scheduling Birth Early on Infant Health. *Journal of the European Economic Association*, 17(1), 30-78.

Brenøe, A. A., & Molitor, R. (2018). Birth order and health of newborns. *Journal of Population Economics*, 31(2), 363-395.

Brunner, B., Kuhn, A. (2014). Announcement effects of health policy reforms: Evidence from the abolition of Austria's baby bonus. *European Journal of Health Economics*, 15, 373-388.

Buckles, K., Guldi, M. (2017). Worth the wait? The effect of early term birth on maternal and infant health. *Journal of Policy Analysis and Management*, 36(4), 748-772.

Chetty, R., Friedman, J., Rockoff, J. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9), 2633-2679.

Clarke, D., Oreffice, S., Quintana-Domeque, C. (2017). On the value of birth weight. HECO WP 2017-018.

Clarke, D., Romano, J., Wolf, M. (2020). The Romano-Wolf multiple hypothesis correction in Stata. *Stata Journal*, 20(4), 812-843.

Conway, K. S., Deb, P. (2005). Is prenatal care really ineffective? Or, is the 'devil'in the distribution? *Journal of Health Economics*, 24(3), 489-513.

Conway, K. S., Kutinova, A. (2006). Maternal health: does prenatal care make a difference? *Health Economics*, 15(5), 461-488.

Corman, H., Dave, D., Reichman, N. E. (2019). The Effects of Prenatal Health on Birth Outcomes: Reconciling a Messy Literature. In *Oxford Research Encyclopedia of Economics and Finance*.

Cunha, F., Heckman, J. (2007). The technology of skill formation. *American Economic Review*, 97(2), 31-47.

Currie, J., Gruber, J. (1996). Saving babies: The efficacy and cost of recent changes in the Medicaid eligibility of pregnant women. *Journal of Political Economy*, *104*(6), 1263-1296.

Currie, J., Grogger, J. (2002). Medicaid expansions and welfare contractions: offsetting effects on prenatal care and infant health? *Journal of Health Economics*, 21(2), 313-335.

Currie, J., Rossin-Slater, M. (2015). Early-lie origins of lifecycle well-being: Research and policy implications. *Journal of Policy Analysis and Management*, 34(1):208-242.

Cygan-Rehm, K. (2016). Parental leave benefit and differential fertility responses: Evidence from a German reform. *Journal of Population Economics*, 29(1), 73-103.

Dench, D., Joyce, T. (2020). The earned income tax credit and infant health revisited. *Health Economics*, 29(1), 72-84.

De Walque, D. (2010). Education, information, and smoking decisions. Evidence from smoking histories in the United States, 1940-2000. *Journal of Human Resources*, 45(3), 682-717.

Diagnoza Społeczna (2009). Warunki i Jakość Życia Polaków. Raport. Czapiński J., Panek T. (Eds.). Rada Monitoringu Społecznego, Warszawa. Available online at http://www.diagnoza.com/pliki/raporty/Diagnoza_raport_2009.pdf [Last accessed: 03.08.2021].

Dickert-Conlin, S., Chandra, A. (1999). Taxes and the timing of births. *Journal of Political Economy*, 107(1), 161-177.

Di Giacomo, Marina, Massimiliano Piacenza, Luigi Siciliani, and Gilberto Turati. "The effect of co-payments on the take-up of prenatal tests." *Journal of Health Economics* 81 (2022): 102553.

Dustmann, C., Schönberg, U. (2012). Expansions in maternity leave coverage and children's long-term outcomes. *American Economic Journal: Applied Economics*, 4(3), 190-224.

Druki Sejmowe (1998). Sprawozdanie Rady Ministrów z realizacji w roku 1997 ustawy z dnia 7 stycznia 1993 r. o planowaniu rodziny, ochronie płodu ludzkiego i warunkach dopuszczalności przerywania ciąży. Archiwum Sejmu (archives of the Polish parliament), Druki Sejmowe III kadencja, Druk nr 592, Warszawa. Available online at http://orka.sejm.gov.pl/Rejestrd.nsf?OpenDatabase [Last accessed: 03.08.2021].

Dziennik Ustaw (2004). Rozporządzenie Ministra Zdrowia z dnia 21 grudnia 2004 r. w sprawie zakresu świadczeń opieki zdrowotnej, w tym badań przesiewowych, oraz okresów, w których te badania są przeprowadzane. Załącznik nr 1: Zakres profilaktycznych świadczeń opieki zdrowotnej u kobiet w okresie ciąży wraz z okresami ich przeprowadzania. Dz. U. 2004 nr. 276 poz. 2740.

East, C. N., Miller, S., Page, M., Wherry, L. R. (2019). Multi-generational Impacts of Childhood Access to the Safety Net: Early Life Exposure to Medicaid and the Next Generation's Health. NBER Working Paper No. 23810. Revised in February 2019.

EURO-PERISTAT (2008). European perinatal health report by the EURO-PERISTAT project in collaboration with SCPE, EUROCAT & EURONEOSTAT. Data from 2004. Coordinated by INSERM, Paris. Available online at https://www.europeristat.com/images/doc/EPHR/european-perinatal-health-report.pdf [Last accessed: 03.08.2021].

EURO-PERISTAT (2013). European perinatal health report. Health and Care of Pregnant Women and Babies in Europe in 2010. Coordinated by INSERM, Paris. Available online at http://www.europeristat.com/reports/european-perinatal-health-report-2010.html [Last accessed: 03.08.2021].

Evans, W., Lien D. (2005). The benefits for prenatal care: Evidence from PAT bus strike. *Journal of Econometrics*, 125, 207-239.

Fertig, A., Watson, T. (2009). Minimum drinking age laws and infant health outcomes. *Journal of Health Economics*, 28(3), 737-747.

Figlio, D., Guryan, J., Karbownik, K., Roth, J. (2014). The effects of poor neonatal health on children's cognitive development. *American Economic Review*, 104(12), 3921-3955.

Firpo, S., Fotin, N., Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3), 953-973.

Frank, R. G., Strobino, D. M., Salkever, D. S., Jackson, C. A. (1992). Updated Estimates of the Impact of Prenatal Care on Birthweight Outcomes By Race. *Journal of Human Resources*, *27*(4), 629-642.

Fryer, R. (2017). The production of human capital in developed countries: Evidence from 196 randomized field experiments. *Handbook of Economic Field Experiments*, 2, 95-322.

Gajate-Garrido, G. (2013). The impact of adequate prenatal care on urban birth outcomes: An analysis in a developing country context. *Economic Development and Cultural Change*, 62(1), 95-130

Gans, J., Leigh, A. (2009). Born on the first of July: An (un)natural experiment in birth timing. *Journal of Public Economics*, 93(1-2), 246-263.

Girouard, N., Imai, Y. (2000). The health care system in Poland. *OECD Economics Department Working Papers*, No. 257.

Godfrey, K., Robinson, S., Braker, D., Osmond, C., Cox, V. (1996). Maternal nutrition in early and late pregnancy in relation to placental and fetal growth. *BMJ*, 312(7028), 410-414.

Gonzalez, F., Kumar, S. (2018). Prenatal care and birthweight in Mexico. *Applied Economics*, 50(10), 1156-1170.

González, L. (2013). The effect of a universal child benefit on conceptions, abortions, and early maternal labor supply. *American Economic Journal: Economic Policy*, 5(3), 160-88.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2), 223-255.

Grossman, M., Joyce, T. J. (1990). Unobservables, pregnancy resolutions, and birth weight production functions in New York City. *Journal of Political Economy*, *98*(5), 983-1007.

Gray, B. (2001). Do Medicaid physician fees for prenatal services affect birth outcomes? *Journal of Health Economics*, *20*(4), 571-590.

Hovdenak, N., Haram, K. (2012). Influence of mineral and vitamin supplements on pregnancy outcome. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 164(2), 127-132.

Hoynes, H., Miller, D., Simon, D. (2015). Income, the earned income tax credit, and infant health. *American Economic Journal: Economic Policy*, 7(1), 172-211.

Huebener, M., Kuehnle, D., Spiess, C. K. (2019). Paid parental leave and child development: Evidence from the 2007 German parental benefit reform and administrative data. *Labour Economics* 61, 101754.

Jones, A. M. (1991). An econometric investigation of low birth weight in the United States. *Journal of Health Economics*, 10(1), 81-99.

Joyce, T. (1999). Impact of augmented prenatal care on birth outcomes of Medicaid recipients in New York City. *Journal of Health Economics*, 18(1), 31-67.

Jukic, A. M., Baird, D. D., Weinberg, C. R., McConnaughey, D. R., Wilcox, A. J. (2013). Length of human pregnancy and contributors to its natural variation. *Human Reproduction*, *28*(10), 2848-2855.

Keskin, P., Shastry, G., Willis, H. (2017). Water quality awareness and breastfeeding: Evidence of health behavior change in Bangladesh. *Review of Economics and Statistics*, 99(2), 265-280.

Kling, J., Liebman, J., Katz, L. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.

Komro, K., Livingston, M., Markowitz, S., Wagenaar, A. (2016). The effect of an increased minimum wage on infant mortality and birth weight. *American Journal of Public Health*, 106(8), 1514-1516.

Kornas-Biela, D. (2012). Male dzieci w Polsce. Zaniedbania prenatalne. *Dziecko krzywdzone. Teoria, badania, praktyka*, 11(11), 24-32.

Lagarde, M., Haines, A., Palmer, N. (2007). Conditional cash transfers for improving uptake of health interventions in low- and middle-income countries. A systematic review. *JAMA*, 298(16); 1900-1910.

Lalive, R., Zweimüller, J. (2009). How does parental leave affect fertility and return to work? Evidence from two natural experiments. *Quarterly Journal of Economics*, 124(3), 1363-1402.

LaLumia, S., Sallee, J., Turner, N. (2015). New evidence on taxes and the timing of birth. *American Economic Journal: Economic Policy*, 7(2), 258-293.

Lange, F. (2011). The role of education in complex health decisions: Evidence from cancer screening. *Journal of Health Economics*, 30(1), 43-54.

Lee, D., Card, D. (2008). Regression discontinuity inference with specification error. Journal of Econometrics, 142(2), 655-674.

Leininger, L., & Levy, H. (2015). Child health and access to medical care. *Future of Children*, *25*(1), 65-90.

Łysoń, P. (2012). Zdrowie i ochrona zdrowia w 2011 R. Informacje i Opracowania Statystyczne, Główny Urząd Statystyczny, Warszawa.

McGeary, K., Dave, D., Lipton, B., Roepr, T. (2020). Impact of comprehensive smoking bans on the health of infants and children. *American Journal of Health Economics*, 6(1), 1-38.

Miller, S., Wherry, L. R. (2019). The long-term effects of early life Medicaid coverage. *Journal of Human Resources*, 54(3), 785-824.

Moller, A. B., Petzold, M., Chou, D., & Say, L. (2017). Early antenatal care visit: a systematic analysis of regional and global levels and trends of coverage from 1990 to 2013. *The Lancet Global Health*, 5(10), 977-983.

Neugart, M., Ohlsson, H. (2013). Economic incentives and the timing of births: Evidence from the German parental benefit reform of 2007. *Journal of Population Economics*, 26, 87-108.

NFZ (2021). Portal Statystyki Narodowego Funduszu Zdrowia (NFZ). Statystyki świadczeń. Available online at: https://statystyki.nfz.gov.pl/ [Last accessed 03.08.2021]

Nieszporska, S. (2017). Priorities in the Polish health care system. *European Journal of Health Economics*, 18(1), 1-5.

Nilsson, P. (2017). Alcohol availability, prenatal conditions, and long-term economic outcomes. *Journal of Political Economy*, 125(4), 1149-1207.

OECD (2021), Caesarean sections (indicator). doi: 10.1787/adc3c39f-en [Last accessed 03.08.2021]

Olafsson, A. (2016). Household financial distress and initial endowments: Evidence from the 2008 financial crisis. *Health Economics*, *25*, 43-56.

Osterman, M. J., Martin, J. A. (2018). Timing and adequacy of prenatal care in the United States, 2016. National Vital Statistics Reports, Volume 67, Number 3, May 2018.

Prina, S., Royer, H. (2014). The importance of parental knowledge: Evidence from weight report cards in Mexico. *Journal of Health Economics*, 37, 232-247.

Rau, T., Sarzosa, M., & Urzúa, S. (2021). The children of the missed pill. *Journal of Health Economics* (forthcoming) doi: 10.1016/j.jhealeco.2021.102496.

Reichman, N. E., Corman, H., Noonan, K., & Dave, D. (2009). Infant health production functions: What a difference the data make. *Health Economics*, *18*(7), 761-782.

Reichman, N., Corman, H., Noonan, K., Schwartz-Soicher, O. (2010). Effects of prenatal care on maternal postpartum behaviors. *Review of Economics of the Household*, 8, 171-197.

Rogala, D., Dylewska, M., Harat, A. (2014). Opieka and kobietą ciężarną w publicznej i prywatnej opiece zdrowotnej. *Pielęgniarstwo Polskie*, 51(1), 13-19.

Romano, J., Wolf, M. (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, 100(469), 94-108.

Romano, J., Wolf, M. (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4), 1237-1282.

Rosenzweig, M. R., Schultz, T. P. (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *Journal of Political Economy*, 91(5), 723-746.

Rossin-Slater, M. (2013). WIC in your neighborhood: New evidence on the impacts of geographic access to clinic. *Journal of Public Economics*, 102, 51-69.

Schulkind, L., Shapiro, T. (2014). What a difference a day makes: Quantifying the effects of birth timing manipulation on infant health. *Journal of Health Economics*, 33, 139-158.

Sonchak, L. (2015). Medicaid reimbursement, prenatal care and infant health. *Journal of Health Economics*, 44, 10-24.

Tamm, M. (2013). The impact of a large parental leave benefit reform on the timing of birth around the day of implementation. *Oxford Bulletin of Economics and Statistics*, 75(4), 585-601.

UNICEF (2021). Maternal, Newborn, Child and Adolescent Health data portal. Available online at: https://www.who.int/data/maternal-newborn-child-adolescent/ [Last accessed 03.08.2021]

UN IGME (2021). United Nations inter-agency group for child mortality estimation. Available online at: https://data.unicef.org/resources/levels-and-trends-in-child-mortality/ [Last accessed 03.08.2021]

Wehby, G., Dave, D., Kaestner, R. (2020). Effects of the minimum wage on infant health. Journal of Policy *Analysis and Management*, 39(2), 411-443.

Wiswall, M., Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *Review of Economic Studies*, 82(2), 791-824.

Wojtyła, C., Wojtyła-Buciora, P. (2016). Polish Pregnancy-related Assessment Monitoring System (Pol-PrAMS): research on lifestyle health behaviours of Polish women during gestation–study design. *Journal of Health Inequalities*, 2(2), 185-191.

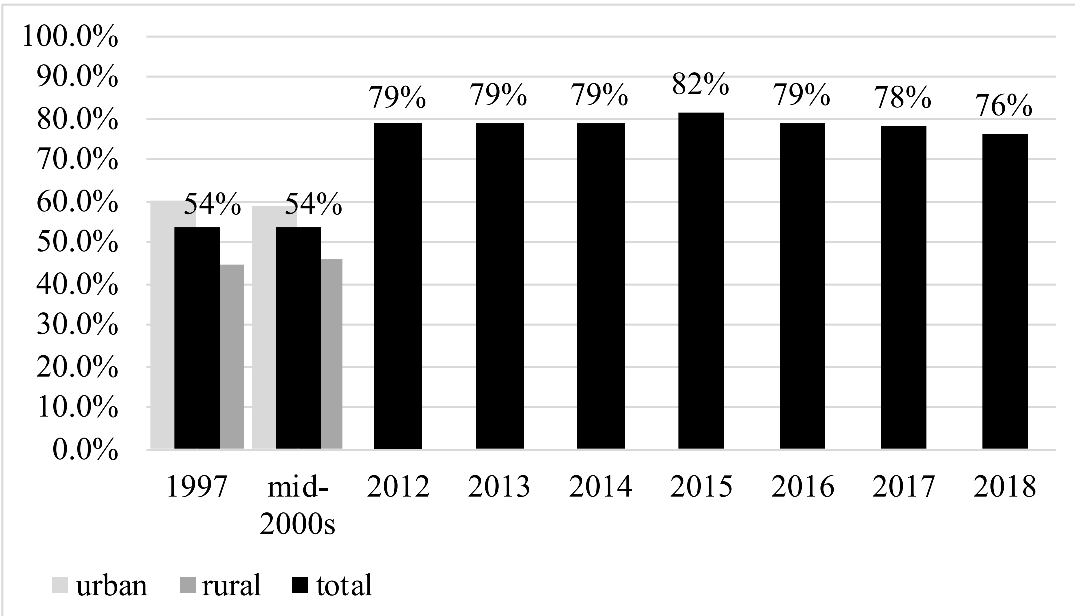
World Health Organization (2016). WHO recommendations on antenatal care for a positive pregnancy experience. WHO Press, Geneva.

Yan, J. (2020). Healthy babies: Does prenatal care really matter? *American Journal of Health Economics*, 6(2), 199-215

Zięba, A. (2006). Podziemie aborcyjne w Polsce. Prezes Zarządu Polskiego Stowarzyszenia Obrońców Życia Człowieka. Wiceprezes Zarządu Polskiej Federacji Ruchów Obrony Życia, Kraków.

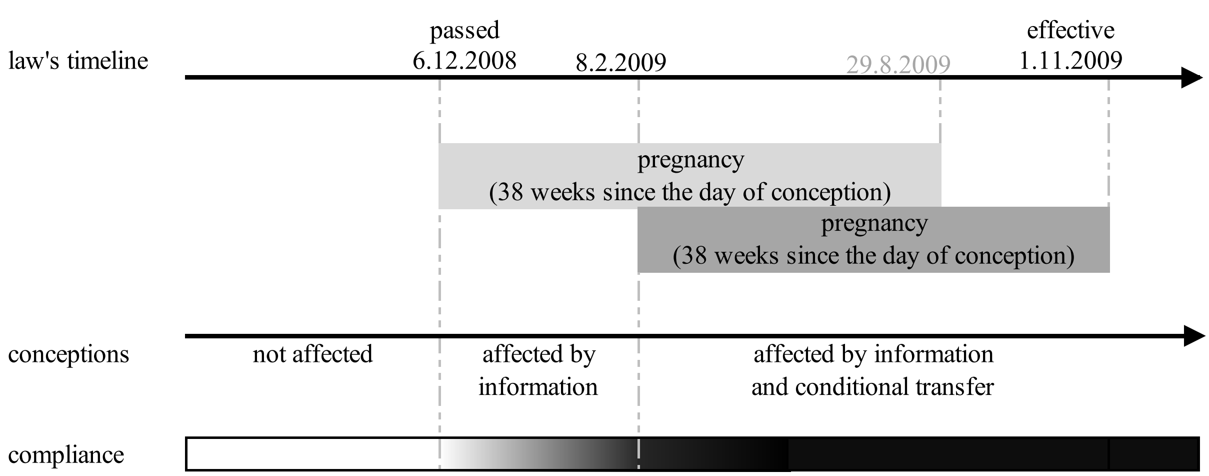
Figures and Tables

Figure 1: Rates of Prenatal Care Initiation in the First Trimester



Note: This figure presents aggregate statistics on rates of prenatal care initiation in the first trimester. The data is limited to visits contracted by the public health care insurance system. The numbers for 1997, are from a governmental report (Druki Sejmowe, 1998); for mid-2000s, from Kornas–Biela (2012). The data from 2012 onwards is provided by the Healthcare Information Systems Center (CSIOZ) of the Ministry of Health.

Figure 2: Timeline of the Law Change



Note: Own illustration. This figure presents the timing of the reform.

Table 1: Descriptive Statistics and Balance of Covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Cutoff: 6th December 2008			Cutoff: 8th February 2009			Donut-hole sample +/- 4 weeks	
	All births	4 weeks before	4 weeks after	D-in-D [p-value]	4 weeks before	4 weeks after	D-in-D [p-value]	Mean	D-in-D [p-value]
				Panel A. Background characteristics					
Age at conception	27.146	26.997	27.049	-0.011 [0.781]	27.989	28.046	0.038 [0.397]	27.268	0.037 [0.371]
First birth	0.511	0.524	0.520	0.003 [0.486]	0.519	0.518	-0.003 [0.453]	0.520	-0.004 [0.181]
Less than high school	0.269	0.250	0.256	0.003 [0.377]	0.253	0.255	0.000 [0.895]	0.274	0.002 [0.511]
High school graduate	0.359	0.359	0.353	-0.010 [0.002]	0.356	0.353	-0.001 [0.880]	0.361	-0.005 [0.119]
In labor force	0.679	0.707	0.704	0.003 [0.371]	0.702	0.703	-0.001 [0.656]	0.678	-0.004 [0.136]
Married	0.799	0.799	0.795	-0.002 [0.491]	0.792	0.789	0.003 [0.389]	0.797	0.000 [0.995]
Divorced	0.027	0.027	0.027	0.000 [0.775]	0.029	0.030	0.001 [0.477]	0.027	0.002 [0.180]
Widowed	0.003	0.003	0.003	0.000 [0.517]	0.003	0.003	0.000 [0.936]	0.003	0.000 [0.837]
Rural residence	0.417	0.411	0.412	-0.002 [0.616]	0.405	0.405	0.001 [0.838]	0.415	0.001 [0.736]
F-statistic				1.430			0.270		0.980
				Panel B. Neonatal outcomes					
Health index (PCA)	0.000	-0.007	0.001	0.007 [0.391]	0.004	0.006	0.008 [0.351]	-0.005	0.020 [0.018]
Fetal death (×100)	0.415	0.432	0.347	-0.089 [0.042]	0.409	0.318	-0.105 [0.038]	0.438	-0.117 [0.022]
Birth weight (grams)	3376.980	3366.124	3368.030	-1.860 [0.671]	3373.216	3373.491	1.415 [0.738]	3367.558	9.728 [0.033]
Low birth weight (×100)	4.844	5.068	4.788	-0.260 [0.083]	4.934	4.920	-0.105 [0.568]	5.053	-0.391 [0.014]
Gestational age (weeks)	39.166	39.179	39.187	0.017 [0.254]	39.195	39.183	0.007 [0.686]	39.184	0.025 [0.130]
Observations	2,327,942	32,639	32,833	313,251	28,394	27,966	284,083	296,828	296,828

Note: Column 1 presents means for the universe of births between 2006 and 2011. Columns 2 and 3 present descriptive statistics for children conceived +/- 4 weeks around 6th December 2008 cutoff while columns 5 and 6 present descriptive statistics for children conceived +/- 4 weeks around 8th February 2009 cutoff. Columns 4 and 7 show the coefficients on an indicator variable that equals to one if a child was conceived in the 4 weeks after the cutoff in the reform year 2008/09 estimated for each listed variable separately as an outcome. The p-value of a significance test is in square brackets and each cell represents a separate linear regression. F-tests in panel A report the value of F-statistics from regressions of the reform indicator on all the covariates listed in this panel. Column 8 presents the mean values for the donut difference-in-differences sample while column 9 presents the balancing test for this sample. Note that some college is reference category in education, not in labor force is a reference category for employment, single mother is a reference category for civil status while urban is a reference category for place of residence. P-values are based on heteroskedasticity robust standard errors.

Table 2: Effects of the Reform on Neonatal Health

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	+/- 1 week		+/- 2 weeks		+/- 3 weeks		+/- 4 weeks	
A. Health index	0.034***	0.035***	0.031***	0.031***	0.021**	0.021**	0.020**	0.021**
	(0.012)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.008)	(0.008)
Pre-reform mean of Y	0.000		0.000		0.000		0.000	
B. Fetal death (×100)	-0.178**	-0.173**	-0.154**	-0.148**	-0.156***	-0.153***	-0.117**	-0.116**
	(0.085)	(0.085)	(0.059)	(0.059)	(0.059)	(0.058)	(0.051)	(0.050)
Pre-reform mean of Y	0.465		0.462		0.450		0.453	
C. Log birth weight	0.767***	0.799***	0.524**	0.539**	0.319	0.332*	0.345**	0.346**
(×100)	(0.212)	(0.212)	(0.214)	(0.218)	(0.196)	(0.201)	(0.166)	(0.169)
Pre-reform mean of Y	810.394		810.431		810.487		810.455	
D. Low birth weight	-0.937***	-0.925***	-0.636***	-0.622***	-0.347*	-0.352*	-0.391**	-0.399**
(×100)	(0.154)	(0.154)	(0.194)	(0.195)	(0.184)	(0.184)	(0.158)	(0.158)
Pre-reform mean of Y	5.099		5.125		5.117		5.129	
E. Gestational age	0.034	0.034	0.042**	0.041**	0.025	0.027	0.025	0.028*
(weeks)	(0.023)	(0.023)	(0.020)	(0.020)	(0.020)	(0.019)	(0.016)	(0.016)
Pre-reform mean of Y	39.180		39.187		39.192		39.189	
N	74,084		148,843		222,848		296,828	
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Each cell is based on a separate regression and displays the coefficient on a Reform dummy from Equation 1. Individual controls include maternal age at conception (linear and squared), dummies for her marital status, education, source of income, province of residence, rural areas, number of previous births, and a child's gender. Standard errors in parentheses are clustered at the level of conception day. The estimation samples include all conceptions up to four weeks before 6th December in years 2005 - 2009 and up to four weeks after 8th February in years 2006 - 2010. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

Table 3: Effects of the Reforms' Information Component on Neonatal Health

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	+/- 1 week		+/- 2 weeks		+/- 3 weeks		+/- 4 weeks	
A. Health index	-0.008	-0.010	-0.007	-0.007	0.002	0.001	0.007	0.007
	(0.012)	(0.012)	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Pre-reform mean of Y	0.000		0.000		0.000		0.000	
B. Fetal death (×100)	-0.084	-0.073	-0.051	-0.048	-0.091*	-0.089*	-0.089**	-0.088**
	(0.058)	(0.057)	(0.059)	(0.059)	(0.050)	(0.050)	(0.044)	(0.044)
Pre-reform mean of Y	0.475		0.457		0.449		0.456	
C. Log birth weight	-0.134	-0.155	-0.201	-0.145	-0.082	-0.048	0.038	0.051
(×100)	(0.217)	(0.217)	(0.171)	(0.170)	(0.173)	(0.175)	(0.152)	(0.151)
Pre-reform mean of Y	810.413		810.416		810.471		810.517	
D. Low birth weight	-0.294	-0.244	-0.068	-0.067	-0.083	-0.083	-0.260*	-0.263*
(×100)	(0.192)	(0.189)	(0.181)	(0.182)	(0.163)	(0.164)	(0.149)	(0.148)
Pre-reform mean of Y	4.990		5.063		5.053		5.026	
E. Gestational age	-0.033	-0.039*	-0.011	-0.015	0.008	0.004	0.017	0.015
(weeks)	(0.020)	(0.021)	(0.018)	(0.018)	(0.017)	(0.017)	(0.015)	(0.015)
Pre-reform mean of Y	39.183		39.186		39.191		39.190	
N	74,796		149,695		228,204		313,251	
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Each cell is based on a separate regression and displays the coefficient on a Reform dummy from Equation 1. The estimation samples include all conceptions up to four weeks before and up to four weeks after 6th December in years 2005 - 2009. Individual controls include maternal age at conception (linear and squared), dummies for her marital status, education, source of income, province of residence, rural areas, number of previous births, and a child's gender. Standard errors in parentheses are clustered at the level of conception day. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

Table 4: Effects of the Reform on Neonatal Health: Sensitivity Analysis

	(1) Health index	(2) Fetal death (×100)	(3) Log birth weight (×100)	(4) Low birth weight (×100)	(5) Gestational age (weeks)
A. Baseline	0.021**	-0.116**	0.346**	-0.399**	0.028*
	(0.008)	(0.050)	(0.169)	(0.158)	(0.016)
Pre-reform mean of Y	0.000	0.453	810.455	5.129	39.189
Observations			296,828		
B. Excluding 2005/06 cohort	0.022**	-0.125**	0.357**	-0.467***	0.030*
	(0.008)	(0.052)	(0.166)	(0.154)	(0.016)
Pre-reform mean of Y	0.000	0.437	810.585	5.071	39.179
Observations			240,365		
C. Excluding 2009/10 cohort	0.017**	-0.118**	0.299*	-0.346**	0.019
	(0.008)	(0.052)	(0.164)	(0.161)	(0.016)
Pre-reform mean of Y	0.000	0.453	810.455	5.129	39.189
Observations			236,006		
D. Lagging control group by 10 weeks	0.020***	-0.130**	0.340**	-0.347**	0.013
	(0.007)	(0.050)	(0.142)	(0.154)	(0.015)
Pre-reform mean of Y	0.000	0.430	810.648	4.985	39.202
Observations			290,769		
E. Assuming mid-week delivery	0.019**	-0.058	0.369**	-0.376**	0.030*
	(0.008)	(0.051)	(0.170)	(0.158)	(0.016)
Pre-reform mean of Y	0.000	0.457	810.466	5.105	39.189
Observations			294,451		
F. Adding quadratic trends	0.024***	-0.108**	0.403**	-0.497***	0.038**
	(0.009)	(0.054)	(0.178)	(0.166)	(0.017)
Pre-reform mean of Y	0.000	0.453	810.455	5.129	39.189
Observations			296,828		
G. Adding maternity leave controls	0.021**	-0.116**	0.346**	-0.399**	0.028*
	(0.008)	(0.050)	(0.169)	(0.158)	(0.016)
Pre-reform mean of Y	0.000	0.453	810.455	5.129	39.189
Observations			296,828		
H. Adding regional unemployment	0.021**	-0.112**	0.388**	-0.405**	0.023
	(0.009)	(0.054)	(0.189)	(0.179)	(0.018)
Pre-reform mean of Y	0.000	0.453	810.455	5.129	39.189
Observations			296,828		
I. Placebo reform (6 months earlier)	0.001	0.019	-0.100	0.109	0.019
	(0.008)	(0.050)	(0.147)	(0.157)	(0.016)
Pre-reform mean of Y	0.000	0.391	811.145	4.688	39.147
Observations			297,435		

Note: Each cell is based on a separate regression and displays the coefficient on a Reform dummy from Equation 1. Samples include all conceptions up to four weeks before 6th December in years 2005 - 2009 and up to four weeks after 8th February in years 2006-2010. All regressions include individual level controls: maternal age at conception (linear and squared), dummies for her marital status, education, source of income, province of residence, rural areas, number of previous births, and a child's gender. Panel B excludes conceptions from 2005/06 cohort while panel C excludes conceptions from 2009/10 cohort. Panel D lags the threshold by 10 weeks to exclude any partially treated women from the control group. Panel E assumes a mid-week delivery in the last (non-completed) gestational week. Panel F includes quadratic trends in conception date. Panel G adds maternity leave controls while panel H adds regional unemployment controls. Panel I presents placebo estimates when we move the two cutoffs back by six months to 6th June 2008 (from 6th December 2008) and 8th August 2008 (from 8th February 2009). Standard errors in parentheses are clustered at the level of conception day. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

Table 5: Exploring Potential Mechanisms: Maternal Health-related Knowledge and Behaviors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Health-related knowledge					Health-related behaviors during and post pregnancy			
	Small alcohol amounts harmful	Large alcohol amounts harmful	Active smoking harmful	Passive smoking harmful	Knows all of this	Alcohol during pregnancy	Smoking during pregnancy	Smoker postpartum	Exposure to smoking at work
Panel A. Baseline June 2009 mean (pre-reform period)									
	0.877	0.946	0.969	0.924	0.832	0.129	0.173	0.091	0.114
Panel B. Unconditional differences between 2009 and subsequent years									
Post 2010	0.096***	0.050***	0.025***	0.070***	0.138***	-0.062***	-0.066***	-0.027***	0.002
	(0.007)	(0.005)	(0.004)	(0.005)	(0.008)	(0.008)	(0.010)	(0.008)	(0.009)
Post 2011	0.107***	0.052***	0.028***	0.073***	0.150***	-0.054***	-0.041***	-0.026***	-0.004
	(0.007)	(0.004)	(0.003)	(0.005)	(0.008)	(0.009)	(0.010)	(0.008)	(0.010)
Post 2012	0.103***	0.051***	0.027***	0.071***	0.144***	-0.055***	-0.059***	-0.035***	-0.012
	(0.007)	(0.004)	(0.004)	(0.005)	(0.008)	(0.008)	(0.010)	(0.007)	(0.009)
Panel C. Differences between 2009 and subsequent years conditional on covariates									
Post 2010	0.099***	0.052***	0.027***	0.069***	0.138***	-0.054***	-0.071***	-0.031***	0.000
	(0.008)	(0.005)	(0.004)	(0.006)	(0.009)	(0.009)	(0.010)	(0.008)	(0.010)
Post 2011	0.120***	0.060***	0.034***	0.074***	0.155***	-0.035**	-0.058***	-0.038***	-0.010
	(0.011)	(0.007)	(0.006)	(0.009)	(0.013)	(0.014)	(0.016)	(0.012)	(0.015)
Post 2012	0.115***	0.059***	0.034***	0.071***	0.149***	-0.034**	-0.069***	-0.042***	-0.011
	(0.012)	(0.008)	(0.006)	(0.009)	(0.013)	(0.014)	(0.016)	(0.012)	(0.016)
Observations	9,372	9,372	9,372	9,372	9,372	9,372	9,372	9,372	8,540

Note: Panel A presents mean values of outcome variables in the pre-reform period while panels B and C present regression output where treatment variables of interest are indicators for survey years 2010, 2011, and 2012, respectively. In these regressions data from 2009 survey serve as a reference category. Regressions in panel B do not include any additional controls while regressions in panel C control for maternal age at delivery (indicators in three-year intervals), dummies for her marital status, education, province of residence, town size, number of previous births, and a child's gender as well as number of health care providers. Heteroskedasticity robust standard errors in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

Online Appendix - Not for Publication

Additional Information on Internal Validity Checks: Analyses of Shifts in Conceptions and Births

This Appendix discusses the details of various tests that we performed to assess the internal validity of our empirical strategy presented in Section 4. First, we examine whether children born up to four weeks before and after the policy-related cutoffs differ in their sociodemographic backgrounds. Table 1 reports summary statistics and formally tests this balancing assumption. Columns 2 and 3 present sample means for children born up to four weeks before and up to four weeks after the 6th December 2008 cutoff. Columns 5 and 6 show the corresponding sample means for the policy implementation threshold on 8th February 2009. In columns 4 and 7, we formally test whether the covariates are balanced across the cutoffs by estimating difference-in-differences regressions similar to Equation 1 but replacing the outcome variable with a particular maternal characteristic and excluding the vector of control variables X_i. For each regression, we report the point estimates on the reform indicator and the corresponding p-values. Column 9 presents an analogous analysis for our donut-hole difference-in-differences sample that excludes the period between 6th December and 8th February in conception-years 2005/6 to 2009/10 from the analysis. Note that columns 4, 7, and 9 present regression adjusted differences and are thus numerically different from the raw differences that can be computed based on their respective treatment-control dyads in the preceding columns.

Results in panel A reveal that, irrespective of the exact cutoff or samples, the treatment and control groups are balanced on almost all covariates. A single statistically significant difference is that mothers conceiving after 6th December 2008 are slightly less likely to be high school graduates; however, this imbalance is small in magnitude and becomes insignificant if we adjust the p-values for multiple hypothesis testing. Additionally, at the bottom of panel A, we report F-statistics from three F-tests where in each estimation sample, we regress the treatment indicator on a full set of background characteristics and test their joint significance. These F-statistics likewise point to the conclusion that maternal characteristics are not systematically related to the reform indicator.

The descriptive statistics for neonatal outcomes presented in panel B preview the results of our main analysis where we mostly compare children who did not receive any treatment (those conceived before 6th December 2008) with those who were treated by both public information and conditionality of the benefit (those conceived after 8th February 2009). The difference-in-differences estimates in column 9 suggest that both policy components together generated statistically significant gains in infant health.

The balancing tests presented in Table 1 argue against the concern that the reform spawned differential sorting of mothers conceiving across thresholds. Hypothetically though, the law announcement on 6th December 2008 could have sped up the decision to have a child if some prospective mothers wanted to avoid the stricter eligibility rules for the "baby bonus" by bearing a child before the law went into effect. If true, then we would also expect a sudden increase in the number of conceptions right after the announcement date. Such immediate effects are unlikely as a healthy, fertile couple typically needs three to six months to conceive a child when actively trying (e.g., González, 2013), whereas the time window to avoid the stricter eligibility rules was limited to two months (between 6th December 2008 and 8th February 2009). Nevertheless, to mitigate this concern, we test for discontinuous changes in the number of pregnancies initiated around both 6th December 2008 and 8th February 2009 by estimating the following equation

$$C_{trac} = \alpha + \beta Reform_{trac} + l_c + f_{dy} + d_{dw} + d_{dw} \times c + k_r + g_a + \varepsilon_{trac} \qquad \text{(A1)}$$

which mimics specification from Equation 1 but replaces the dependent variable with a (log) number of conceptions occurring on day t, in a region r, age group a, in a given winter (cohort) c. Thus, to estimate this equation, we aggregate our individual-level data to cells by the inferred day of conception, maternal region of residence (being an interaction between province and a dummy for rural areas), and maternal age at conception. In these regressions, we also include region (k_r), and age (g_a) fixed effects. In this case, $Reform_{trac}$ is an indicator equal to one for conceptions taking place on or after 6th December 2008 (or 8th February 2009), and β captures any excess daily conceptions occurring shortly after compared to shortly before these dates net of any date-specific conception patterns differenced out by comparison with surrounding years. If the reform announcement accelerated conceptions, we expect β to be positive and statistically significant.

Panels A and B in Table A2 present the results of this analysis for four different samples where the time span around the cutoff varies between one and four weeks. Irrespective of the exact sample chosen, the estimates corresponding to 6th December 2008 cutoff imply between 0.002 and 0.027 fewer conceptions per day, on average, in days following the reform announcement. This effect is not only statistically insignificant but also small in magnitude given the sample means of over 2 conceptions per cell. This is confirmed by the specifications using the log number of conceptions as an outcome, which imply effect sizes between 0 and 0.8 percent. Panel B shows that there are also no statistically significant or sizable changes in the number of conceptions after 8th February 2009. Note that any increases in the number of conceptions could be not only due to endogenous delays but also due to early prenatal care improving the survivability of fetuses beyond the 22nd week of gestation, which is the relevant threshold for filing a birth declaration (i.e., being recorded in the data). Fewer early deaths of the "marginal" fetuses would bias our main results in Section 5 downwards because of the negative compositional effect on health outcomes.

Another potential composition effect could emerge if mothers with due dates around 1st November 2009 sped up a delivery (e.g., by labor induction or caesarian section) to apply for the "baby bonus" without the additional requirement mandated by the new law. We test for such endogenous shifting of births by estimating the following equation:

$$B_{trac} = \alpha + \beta Reform_{trac} + l_c + f_{dy} + d_{dw} + d_{dw} \times c + k_r + g_a + \varepsilon_{trac} \qquad \text{(A2)}$$

which is very similar to Equation A1, but the dependent variable here is a (log) number of births occurring on day t, in region r, age group a, in a given winter (cohort) c. Thus, for estimations, we aggregate the data to cells using the day of birth, region (being an interaction between province and a dummy for rural areas), and maternal age at birth. In this case, $Reform_{trac}$ is an indicator equal to one for births taking place on or after 1st November 2009, and thus β captures any changes in the number of deliveries occurring right after compared to right before the reform's implementation day and net of any date-specific birth patterns differenced out by comparison with surrounding years. If the reform's implementation endogenously led to speeding up deliveries, we would expect β to be negative and statistically significant.

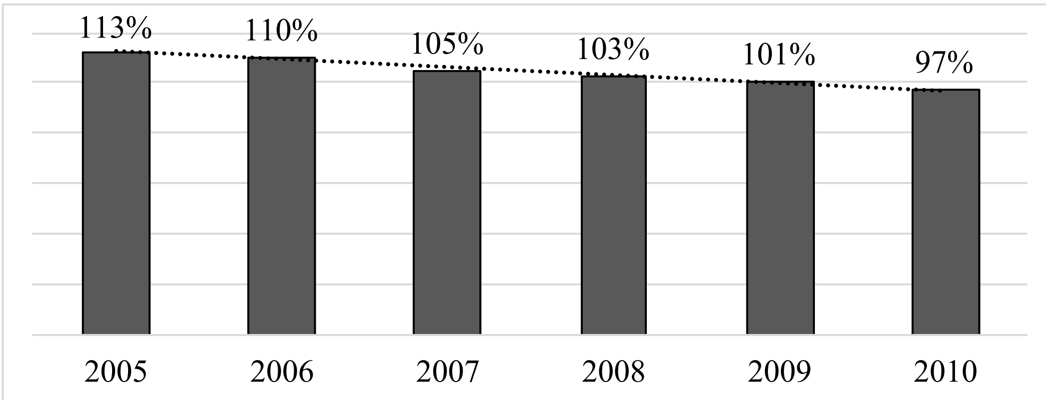
Panel C in Table A2 presents the regression results for four bandwidths around the 1st November 2009 cutoff. The point estimates imply statistically significant increases in mass of deliveries right after the new law became effective, which contradicts the prediction outlined above. Instead, these results suggest that the reform could have prolonged the average duration of pregnancies, which we view as a positive outcome. In terms of magnitude, the increase in births by 2 to 4 percent is relatively small compared to the shifts in births due to financial incentives reported by Borra et al. (2019) who found effect sizes in the range of 11 to 38 percent. Our findings go also in the opposite direction as Borra et al. (2019) show that revoking a "baby bonus" leads mothers to speed a birth up (through induction or c-section), so that they still qualify for the payment. In contrast, in our case, the additional constraint on the eligibility for a financial benefit leads to small increases in births after it becomes enforceable. Thus, our finding is consistent with positive effects of the intervention on pregnancy length rather than with manipulating birth date through e.g., a c-section. In 2009, 23% of births in Poland were delivered by a c-section, which was comparable to the level in UK (24%) and below the level in the US (33%) and many European countries (e.g., 30% in Germany). Since then, however, the c-section rate in Poland increased dramatically to over 40% in most recent years (OECD 2019). Unfortunately, our data do not have information on method of delivery.

We also estimated all Panels in Table A2 at a more aggregate level, i.e., solely by conception day (Panels A and B) or birth day (Panel C) and found quantitatively similar results. However, the effects were less precisely estimated due to the smaller number of cells (between 40 in column 1 and 280 in column 4). Thus, we prefer the results presented here as more disaggregated data increase the statistical power.

Finally, our DD approach also requires that, absent the policy change, conceptions and births around the cutoffs would have followed similar seasonality patterns as their counterparts for the nonreform years. Only under this assumption, we can net out seasonality using double differencing. We present graphical evidence in favor of this assumption in Figure A7 in the Online Appendix, which shows that the number of conceptions and births in reform years tracks very closely with those in the surrounding years. Although there are clear seasonality patterns for all years – particularly related to Christmas holidays and New Year's Day – we do not observe any discontinuities across years around our cutoffs of interest.

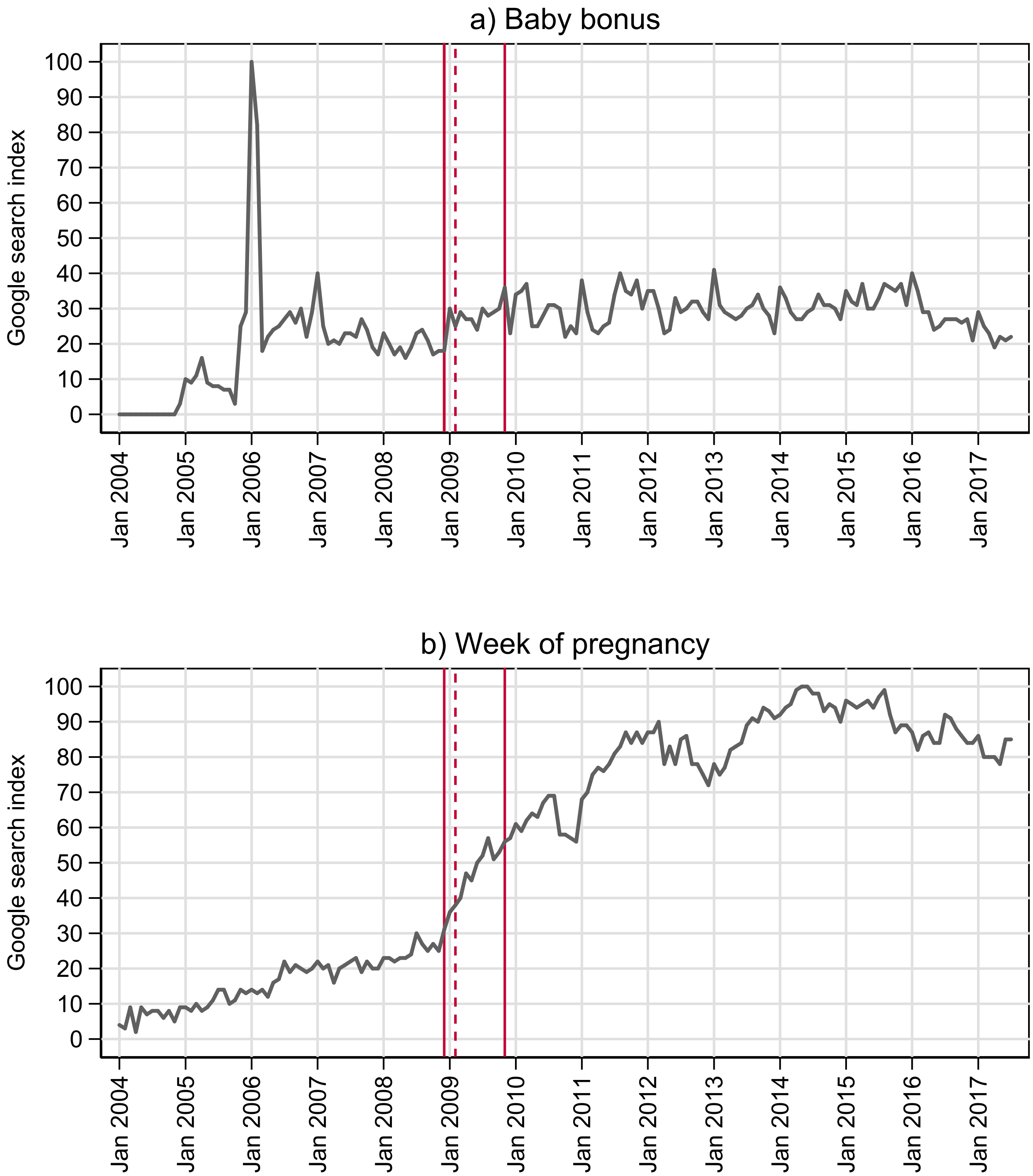
Additional Figures

Figure A1: Obstetrics and Gynecology Visits with the Public Health Care System Providers Among Pregnant Women



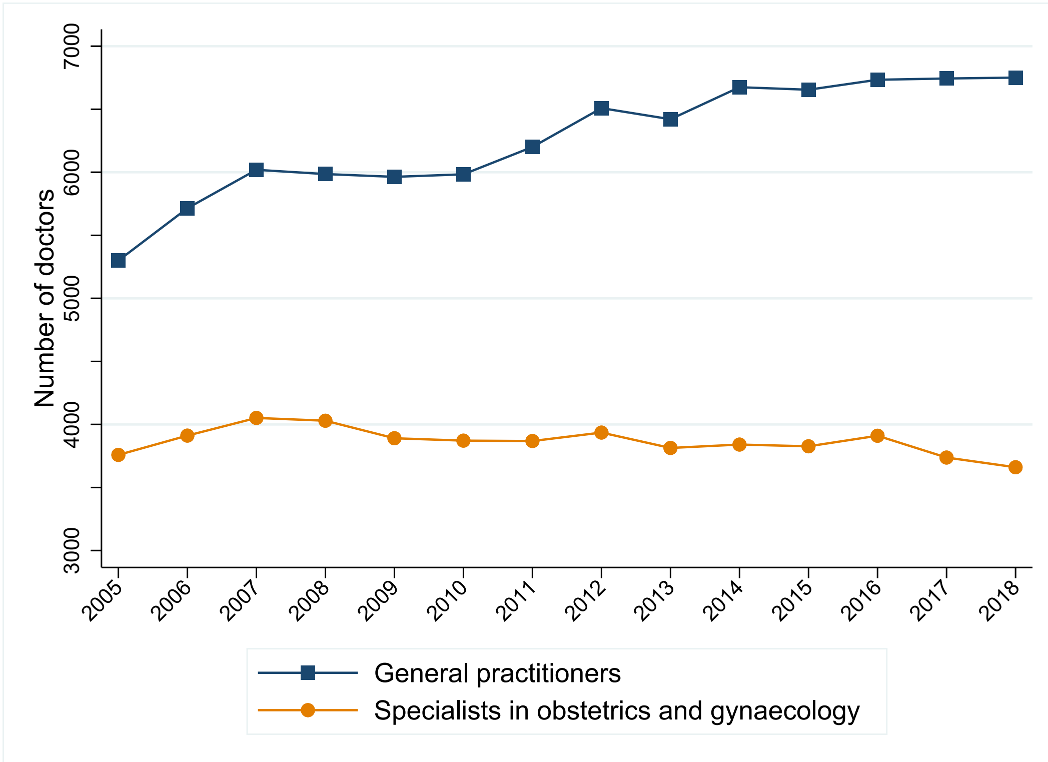
Note: The figure relates the number of women with a diagnosed pregnancy who generated at least one bill for outpatient specialist care in obstetrics/gynecology within the public health care system to the annual number of births. Diagnosed pregnancies refer to ICD-10 codes O00-O99 and Z32-Z36 and include pregnancies ended by a miscarriage or abortion (approximately 11% of all births as shown in Figure A6). The negative trend likely reflects the steadily increasing coverage by private add-on insurances. In 2009, 10 percent of households had an add-on private insurance. Source of data: National Health Fund of Poland (Narodowy Fundusz Zdrowia - NFZ).

Figure A2: Internet Searches for "Baby Bonus" and "Week of Pregnancy"



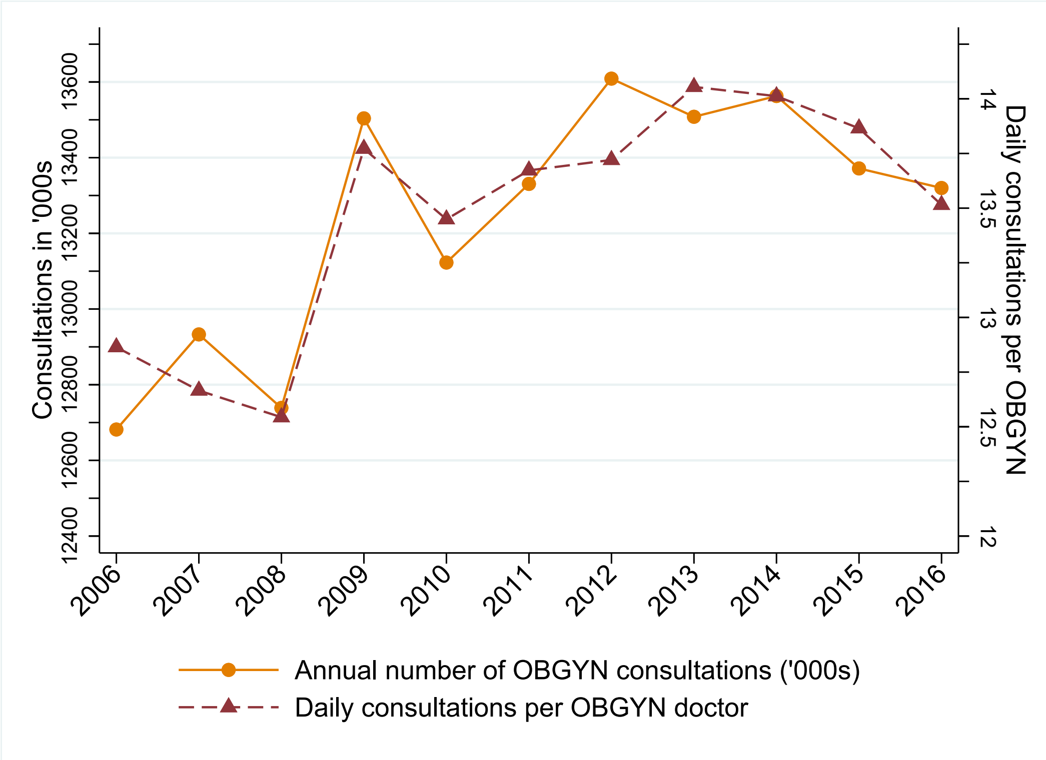
Note: Google Trends available online at https://trends.google.com/trends/explore [Last accessed: 01.08.2019]. The figures show monthly Google search index for the Polish keywords "becikowe" and "tydzień ciąży", respectively. The search region was restricted to Poland and the search period to 01.01.04 - 31.07.17. The vertical solid lines mark December 2008 and November 2009, respectively. The vertical dashed line marks February 2009.

Figure A3: Number of General Practitioners as well as Gynaecologists and Obstetricians



Note: This figure presents aggregate statistics on number of general practitioners as well as OB/GYN specialists in Poland between 2005 and 2018. This data is limited to doctors under contract to the public health care insurance system.
Source: The Central Statistics Office of Poland, Basic Data on Health Care (various years).

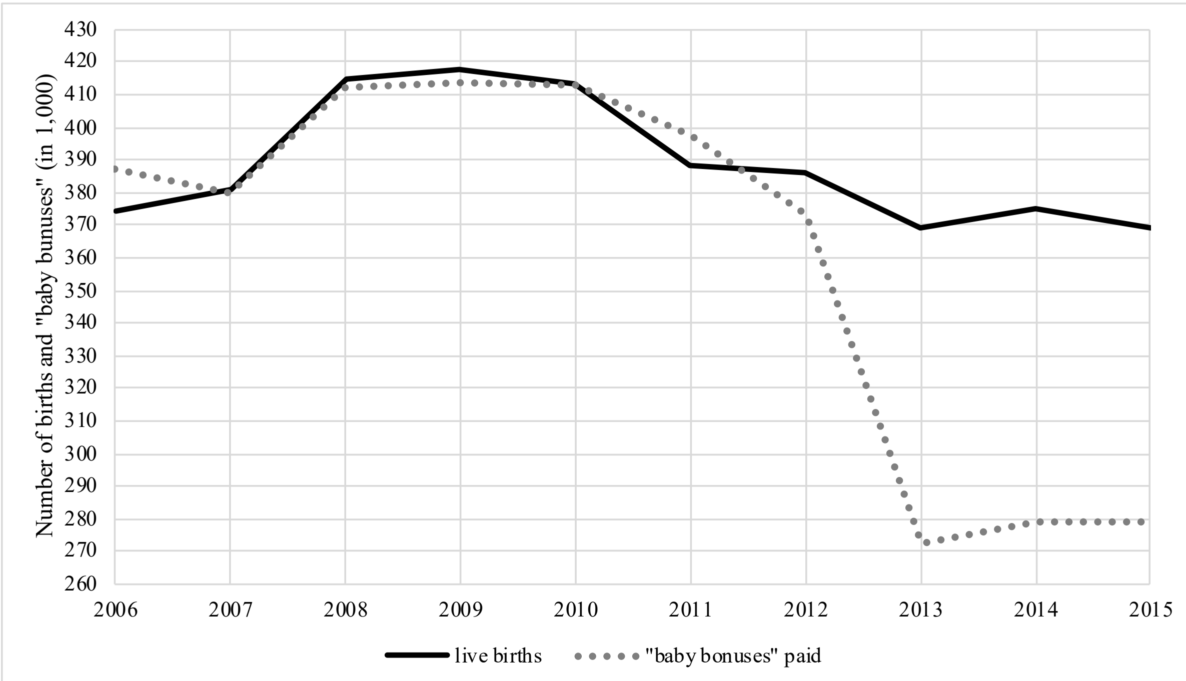
Figure A4: Gynaecologists and Obstetricians Consultations



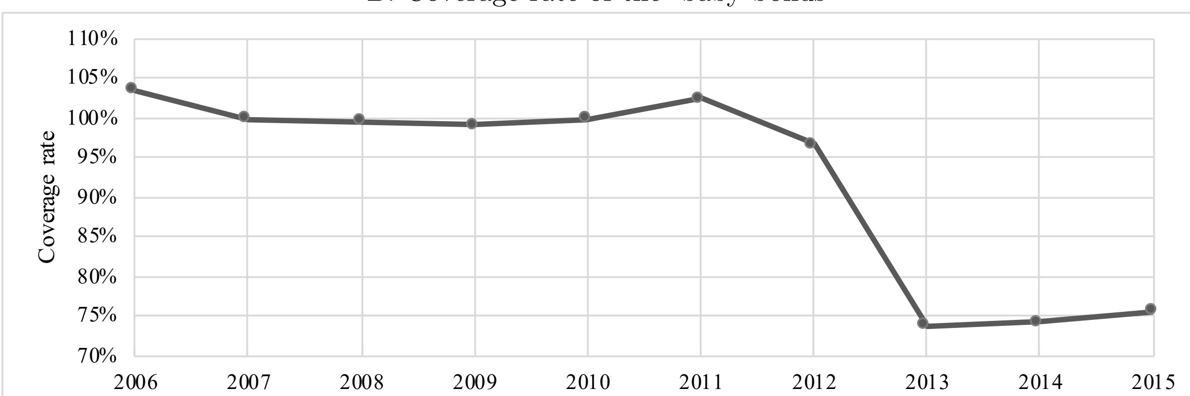
Note: This figure presents aggregate statistics on the annual number of consultations given by OB/GYN specialists in Poland between 2006 and 2016. This data is limited to doctors under contract to the public health care insurance system.
Source: The Central Statistics Office of Poland, Basic Data on Health Care (various years).

Figure A5: Take up of the "Baby Bonus" Between 2006 and 2015

A. Absolute number of births and "baby bonuses" paid



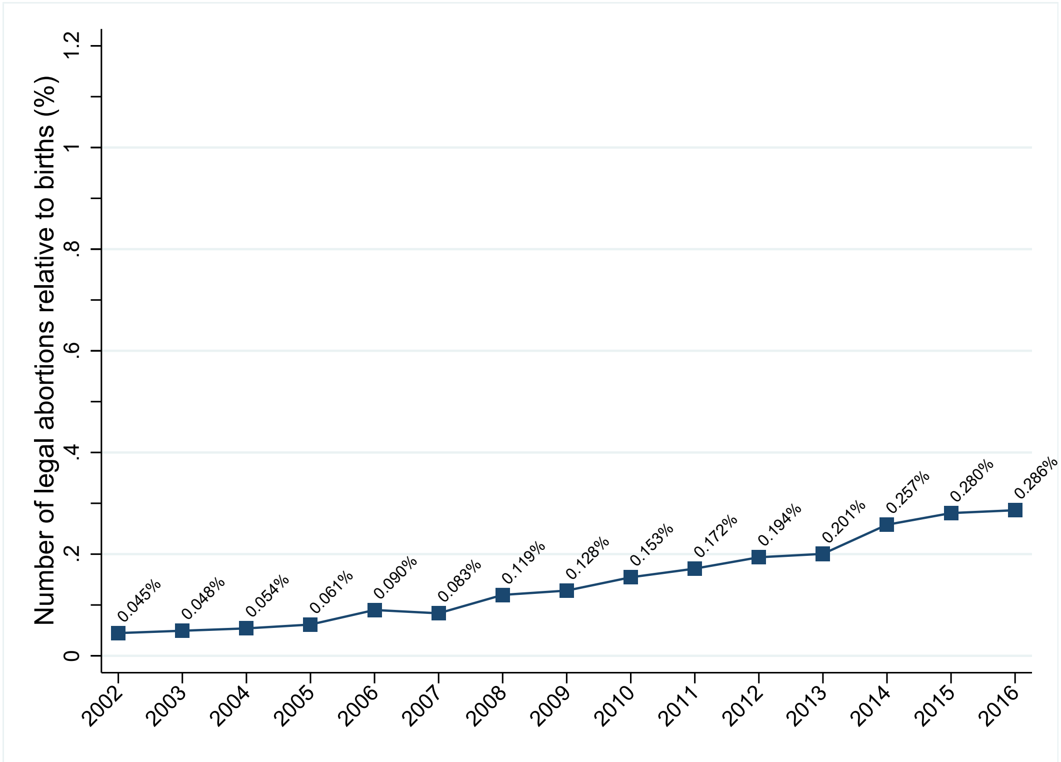
B. Coverage rate of the "baby bonus"



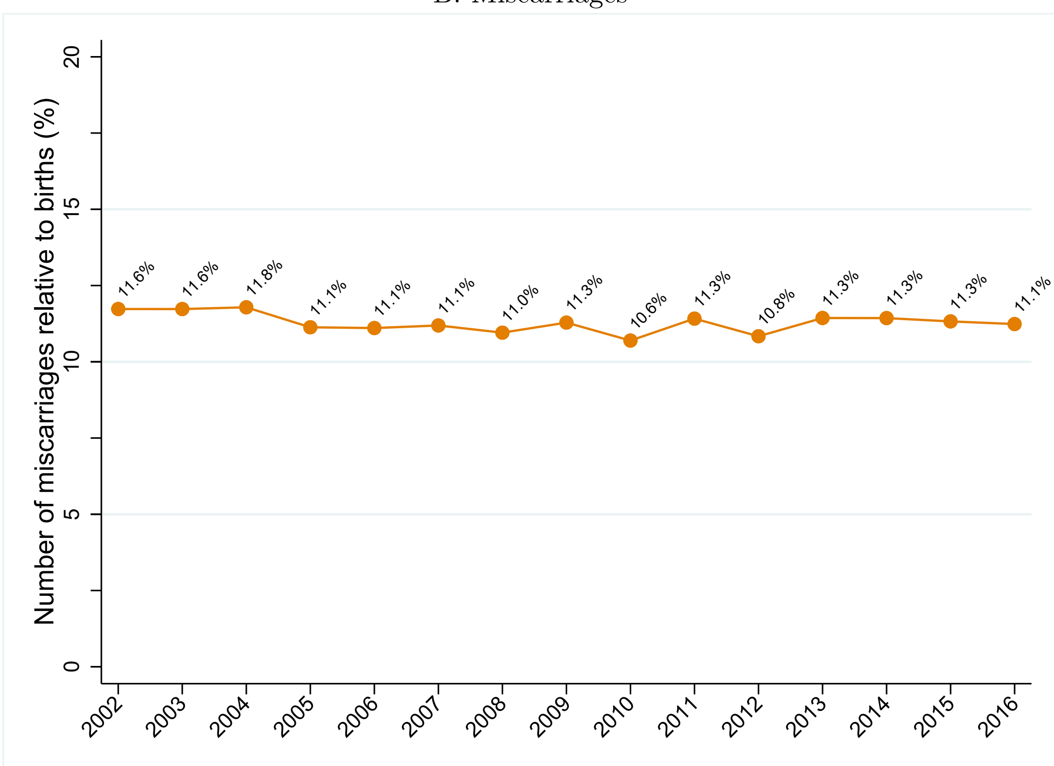
Note: This figure presents aggregate statistics on all births and number of benefits paid based on annual data from The Ministry of Family, Labour and Social Policy. The coverage rate is a ratio of the number of "baby bonuses" to the number of all births in a given year. In 2006, the coverage rate exceeded 100% because upon its introduction, the benefit was paid retrospectively to parents of children born after 9th November 2005. Likewise, the excess rate in 2011 reflects the retrospective payments to mothers who initially lost eligibility in the preceding years and after relaxing the criteria in 2010 could re-apply for the benefit. The sharp decline in utilization observed between 2012 and 2013 reflects an introduction of income eligibility criteria for "baby bonus", thus changing it from universal to means tested.

Figure A6: Legal Abortions and Miscarriages between 2002 and 2016 in Poland

A. Legal abortions

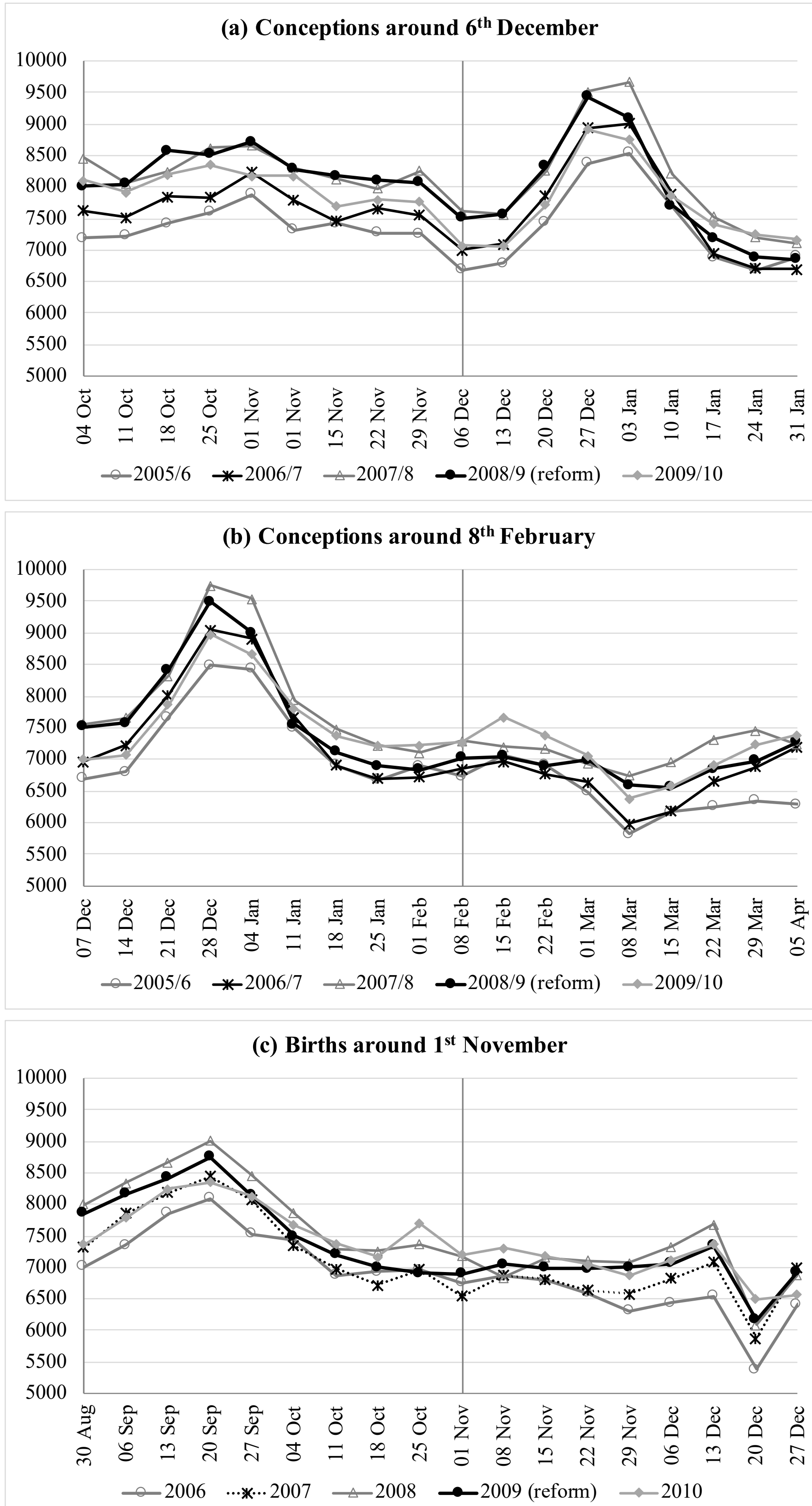


B. Miscarriages



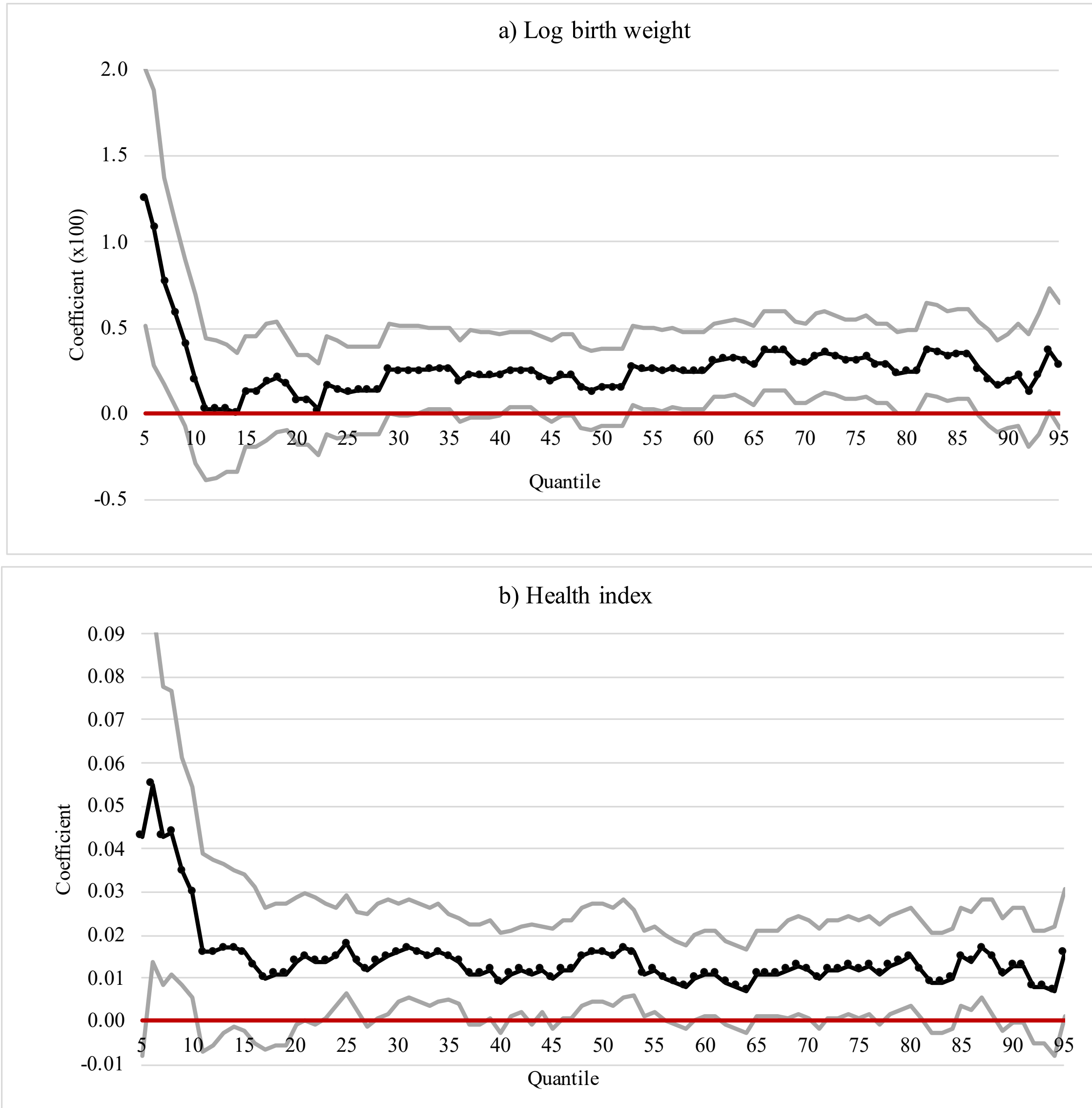
Note: This figure presents aggregate statistics on number of legal abortions and miscarriages relative to live births between 2002 and 2016 in Poland.
Source: Centrum Systemow Informacyjnych Ochrony Zdrowia (CSIOZ).

Figure A7: Trends in Conceptions and Births



Note: This figure plots the weekly number of conceptions and births across years around three cutoffs: 6th of December, 8th of February, and 1st of November. Conception date is calculated as the day of birth minus the actual completed gestational weeks converted into days. The x-axis shows the starting days of seven-day periods.

Figure A8: Effects of the Reform on Neonatal Health: Quantile Regressions



Note: The estimation samples include all conceptions up to four weeks before 6th December in years 2005 - 2009 and up to four weeks after 8th February in years 2006 - 2010. These figures plot unconditional quantile regression estimates, from RIF regressions (Firpo et al., 2009) implemented using the rifreg command in Stata, on the reform dummy as defined in Equation 1. Solid black line depicts the point estimate while gray lines present 90% confidence intervals. Panel A presents estimates for log birth weight while panel B for health index. Control variables include maternal age at delivery (linear and squared), dummies for her marital status, education, source of income, province of residence, rural areas, number of previous births, and a child's gender.

Additional Tables

Table A1: Construction of Principal Components Health Index

	First component	Second component
Fetal death (indicator)	-0.389	0.918
Birth weight (grams)	0.643	0.328
Gestational age (weeks)	0.660	0.221
Eigenvalue	1.708	0.877
Summary statistics for the first component		
Mean	0.000	
Standard deviation	1.307	

Note: This table reports the results of a principal components analysis of an indicator for fetal death, birth weight in grams, and gestational age in weeks. The eigenvectors associates with the first and second components are reported, as well as their associated eigenvalues. The bottom panel reports summary statistics of the health index, defined as the first component of the principal components analysis, for the overall sample.

Table A2: Effects of the Reform Passage and Implementation on Number of Conceptions and Births

	(1) +/- 1 week	(2) +/- 2 weeks	(3) +/- 3 weeks	(4) +/- 4 weeks
	Panel A. 6th December 2008 cutoff (conceptions)			
Number of conceptions	-0.002	-0.027	-0.025	-0.017
	(0.039)	(0.028)	(0.023)	(0.020)
Pre-reform mean of Y	2.271	2.268	2.286	2.313
Log number of conceptions	0.000	-0.008	-0.008	-0.008
	(0.014)	(0.010)	(0.008)	(0.007)
Observations	33,004	66,072	99,811	134,962
	Panel B. 8th February 2009 cutoff (conceptions)			
Number of conceptions	0.052	0.006	-0.007	0.006
	(0.037)	(0.026)	(0.021)	(0.019)
Pre-reform mean of Y	2.149	2.157	2.164	2.178
Log number of conceptions	0.014	0.001	-0.003	0.000
	(0.014)	(0.010)	(0.008)	(0.007)
Observations	32,364	64,707	97,147	129,977
	Panel C. 1st November 2009 cutoff (births)			
Number of births	0.094**	0.089***	0.038*	0.043**
	(0.037)	(0.026)	(0.021)	(0.019)
Pre-reform mean of Y	2.165	2.159	2.163	2.174
Log number of births	0.038***	0.039***	0.020**	0.022***
	(0.014)	(0.010)	(0.008)	(0.007)
Observations	32,365	64,626	97,103	129,855

Note: Panels A and B show the coefficient on a Reform dummy from Equation A1 obtained for the cutoffs of 6th December and 8th February, respectively. Panel C shows the coefficient on a Reform dummy from Equation A2. The dependent variable is the (log) daily number of conceptions (Panels A and B) or births (Panel C) at day-by-province-by-rural-by-mother's age cells. Each cell represents a separate linear regression. Heteroskedasticity robust standard errors in parentheses. The estimation samples include all conceptions/births up to four weeks before and up to four weeks after a particular cutoff in years 2005/6 - 2009/10. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

Table A3: Heterogeneity in the Effects of the Reform on Neonatal Health

	(1) Health index	(2) Fetal death (×100)	(3) Log birth weight (×100)	(4) Low birth weight (×100)	(5) Gestational age (weeks)
		Panel A: Baseline specification			
Reform	0.020**	-0.117**	0.345**	-0.391**	0.025
	(0.008)	(0.051)	(0.166)	(0.158)	(0.016)
Pre-reform mean of Y	0.000	0.453	810.455	5.129	39.189
		Panel B: Interaction with education			
High school dropout	0.040***	-0.234***	0.636**	-0.617	0.049
(reference)	(0.015)	(0.078)	(0.286)	(0.380)	(0.032)
Reform × High school	-0.027	0.183*	-0.450	0.497	-0.031
graduate	(0.019)	(0.109)	(0.375)	(0.496)	(0.040)
Reform × Some	-0.027*	0.136*	-0.334	0.108	-0.023
college and above	(0.016)	(0.080)	(0.317)	(0.403)	(0.035)
		Panel C: Interaction with labor force participation			
Not in labor force	0.025*	-0.126*	0.426*	-0.702**	0.038
(reference)	(0.014)	(0.071)	(0.253)	(0.284)	(0.031)
Reform × Working	-0.006	0.014	-0.114	0.431	-0.013
	(0.014)	(0.074)	(0.255)	(0.314)	(0.033)
		Panel D: Interaction with place of residence			
Rural (reference)	0.025**	-0.115*	0.479**	-0.428*	0.022
	(0.012)	(0.067)	(0.229)	(0.227)	(0.024)
Reform × Urban	-0.006	-0.002	-0.223	0.049	0.011
	(0.011)	(0.065)	(0.234)	(0.272)	(0.027)
		Panel E: Interaction with maternal age at conception			
Teen birth (reference)	0.056*	-0.239*	1.272*	-2.407***	0.074
	(0.032)	(0.132)	(0.654)	(0.700)	(0.071)
Reform × Non-teen	-0.037	0.130	-0.974	2.113***	-0.048
birth	(0.033)	(0.140)	(0.656)	(0.690)	(0.071)
		Panel F: Interaction with birth order			
First birth (reference)	0.014	-0.069	0.259	-0.351*	0.018
	(0.011)	(0.069)	(0.211)	(0.211)	(0.020)
Reform × Non-first	0.014	-0.098	0.180	-0.099	0.021
birth	(0.013)	(0.081)	(0.249)	(0.252)	(0.028)
N			296,828		

Note: Each column in each panel is based on a separate regression without controls based on the donut-hole sample where we include conceptions up to four weeks before 6th December in years 2005-2009 and up to four weeks after 8th February in years 2006-2010. It displays the coefficient on a Reform dummy from Equation 1, which is labelled as reference category group comparison between treated and control individuals, as well as its interactions with demographic characteristics. Standard errors in parentheses are clustered at the level of conception day. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

Table A4: Robustness: Event Study Analysis

	(1) Health index	(2) Fetal death (×100)	(3) Log birth weight (×100)	(4) Low birth weight (×100)	(5) Gestational age (weeks)
13-15 weeks prior to reform	-0.014	0.024	-0.239	0.073	-0.013
	(0.010)	(0.064)	(0.174)	(0.176)	(0.018)
10-12 weeks prior to reform	-0.001	-0.063	-0.106	-0.048	0.003
	(0.010)	(0.057)	(0.198)	(0.178)	(0.017)
7-9 weeks prior to reform	0.009	-0.072	0.198	-0.346*	0.006
	(0.011)	(0.052)	(0.203)	(0.181)	(0.019)
4-6 weeks prior to reform	-0.003	-0.046	0.011	0.015	-0.012
	(0.010)	(0.060)	(0.193)	(0.189)	(0.018)
1-3 weeks prior to reform			Reference period		
1-3 weeks after reform signing	0.003	-0.092	-0.013	-0.105	0.007
	(0.010)	(0.056)	(0.183)	(0.174)	(0.018)
4-6 weeks after reform signing	0.016*	-0.096*	0.205	-0.141	0.034*
	(0.010)	(0.051)	(0.193)	(0.204)	(0.018)
7-9 weeks after reform signing	0.017	-0.024	0.411*	-0.350*	0.018
	(0.011)	(0.061)	(0.220)	(0.214)	(0.019)
1-3 weeks after reform implementation	0.032***	-0.202***	0.510**	-0.491**	0.041**
	(0.011)	(0.059)	(0.211)	(0.205)	(0.020)
4-6 weeks after reform implementation	0.022***	-0.015	0.352*	-0.258	0.055***
	(0.010)	(0.065)	(0.209)	(0.203)	(0.018)
7-9 weeks after reform implementation	0.019*	0.003	0.356*	-0.421**	0.043**
	(0.011)	(0.057)	(0.210)	(0.201)	(0.020)
10-12 weeks after reform implementation	0.013	0.006	0.297	-0.476**	0.030
	(0.012)	(0.071)	(0.222)	(0.193)	(0.021)
13-15 weeks after reform implementation	0.019*	0.046	0.545***	-0.456**	0.035*
	(0.010)	(0.072)	(0.204)	(0.200)	(0.019)
Reference period mean	0.000	0.431	810.645	5.014	39.180
N			1,459,814		

Note: Each columns is based on a separate regression with controls and the sample includes conceptions up to 15 weeks before and up to 24 weeks after 6th December in years 2005-2010. Control variables are identical to those used in Equation 1. The 24-week post-period is divided into 9-week period between signing and implementation of the reform and 15-week period post implementation. Event-time dummies are defined based on 3-week intervals and a period of 1 to 3 weeks prior to reform signing serves as a reference period. Standard errors in parentheses are clustered at the level of conception day. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

Table A5: Effects of the Reform on Neonatal Health: Adjusting for Multiple Hypotheses Testing and Alternative Clustering

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	+/- 1 week		+/- 2 weeks		+/- 3 weeks		+/- 4 weeks	
A. Health index	**0.034**	**0.035**	**0.031**	**0.031**	**0.021**	**0.021**	**0.020**	**0.021**
Model p-value	0.004	0.003	0.003	0.003	0.047	0.039	0.018	0.013
Resample p-value	0.020	0.016	0.004	0.004	0.012	0.012	0.004	0.004
Romano-Wolf p-value	0.028	0.024	0.012	0.012	0.068	0.036	0.020	0.024
Month-level clustering p-value	0.006	0.005	0.001	0.001	0.001	p<0.001	p<0.001	p<0.001
B. Fetal death (×100)	**-0.178**	**-0.173**	**-0.154**	**-0.148**	**-0.156**	**-0.153**	**-0.117**	**-0.116**
Model p-value	0.039	0.046	0.010	0.013	0.008	0.009	0.022	0.022
Resample p-value	0.068	0.076	0.012	0.016	0.008	0.008	0.016	0.016
Romano-Wolf p-value	0.147	0.143	0.020	0.020	0.012	0.016	0.032	0.024
Month-level clustering p-value	0.003	0.004	0.001	0.002	p<0.001	p<0.001	p<0.001	p<0.001
C. Log birth weight (×100)	**0.767**	**0.799**	**0.524**	**0.539**	**0.319**	**0.332**	**0.345**	**0.346**
Model p-value	0.001	p<0.001	0.015	0.015	0.106	0.100	0.039	0.042
Resample p-value	0.020	0.012	0.008	0.008	0.056	0.048	0.004	0.004
Romano-Wolf p-value	0.020	0.016	0.020	0.020	0.080	0.076	0.032	0.024
Month-level clustering p-value	0.001	0.001	0.003	0.001	0.006	0.003	p<0.001	p<0.001
D. Low birth weight (×100)	**-0.937**	**-0.925**	**-0.636**	**-0.622**	**-0.347**	**-0.352**	**-0.391**	**-0.399**
Model p-value	p<0.001	p<0.001	0.001	0.002	0.061	0.057	0.014	0.012
Resample p-value	0.004	0.004	0.008	0.008	0.052	0.048	0.004	0.012
Romano-Wolf p-value	0.004	0.004	0.012	0.012	0.080	0.076	0.016	0.024
Robust EHW errors p-value	p<0.001	p<0.001	0.001	0.001	0.025	0.020	0.002	0.001
E. Gestational age (weeks)	**0.034**	**0.034**	**0.042**	**0.041**	**0.025**	**0.027**	**0.025**	**0.028**
Model p-value	0.137	0.146	0.037	0.041	0.203	0.162	0.130	0.078
Resample p-value	0.199	0.191	0.028	0.020	0.116	0.064	0.044	0.040
Romano-Wolf p-value	0.199	0.191	0.028	0.020	0.116	0.076	0.044	0.040
Month-level clustering p-value	0.066	0.062	0.009	0.007	0.012	0.004	0.005	0.001
N	74,084		148,843		222,848		296,828	
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: This table replicates results from Table 2 adjusting for multiple hypotheses testing and using standard errors clustered at conception month level. Multiple hypothesis adjustment uses procedure outlined in Romano and Wolf (2005a) and Romano and Wolf (2005b) and programmed by Clarke et al. (2020). Each estimate, in bold, is based on a separate regression and displays the coefficient on a Reform dummy from Equation 1. We then provide p-value from standard model as in Table 2 followed by resampled p-values, by Romano-Wolf adjusted p-values, and by p-values based on standard errors clustered at month of conception level. When using Clarke et al. (2020) procedure we bootstrap 250 times.

Table A6: Effects of the Reforms' Information Component on Neonatal Health: Adjusting for Multiple Hypotheses Testing and Alternative Clustering

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	+/- 1 week		+/- 2 weeks		+/- 3 weeks		+/- 4 weeks	
A. Health index	**-0.008**	**-0.010**	**-0.007**	**-0.007**	**0.002**	**0.001**	**0.007**	**0.007**
Model p-value	0.537	0.411	0.450	0.468	0.853	0.867	0.391	0.387
Resample p-value	0.594	0.494	0.438	0.462	0.785	0.785	0.315	0.283
Romano-Wolf p-value	0.729	0.649	0.737	0.781	0.928	0.936	0.375	0.390
Month-level clustering p-value	0.541	0.439	0.377	0.421	0.739	0.778	0.095	0.104
B. Fetal death (×100)	**-0.084**	**-0.073**	**-0.051**	**-0.048**	**-0.091**	**-0.089**	**-0.089**	**-0.088**
Model p-value	0.148	0.205	0.388	0.420	0.068	0.075	0.042	0.044
Resample p-value	0.219	0.291	0.406	0.426	0.032	0.036	0.016	0.016
Romano-Wolf p-value	0.514	0.649	0.737	0.781	0.155	0.171	0.088	0.088
Month-level clustering p-value	0.002	0.005	0.183	0.219	0.007	0.008	0.001	0.001
C. Log birth weight (×100)	**-0.134**	**-0.155**	**-0.201**	**-0.145**	**-0.082**	**-0.048**	**0.038**	**0.051**
Model p-value	0.538	0.478	0.244	0.396	0.635	0.784	0.803	0.735
Resample p-value	0.630	0.582	0.271	0.414	0.578	0.741	0.789	0.733
Romano-Wolf p-value	0.729	0.649	0.614	0.781	0.928	0.936	0.789	0.733
Month-level clustering p-value	0.574	0.544	0.173	0.322	0.514	0.708	0.699	0.605
D. Low birth weight (×100)	**-0.294**	**-0.244**	**-0.068**	**-0.067**	**-0.083**	**-0.083**	**-0.260**	**-0.263**
Model p-value	0.130	0.201	0.707	0.714	0.612	0.611	0.083	0.077
Resample p-value	0.247	0.303	0.729	0.749	0.657	0.618	0.068	0.064
Romano-Wolf p-value	0.514	0.649	0.785	0.781	0.928	0.908	0.116	0.116
Month-level clustering p-value	0.094	0.184	0.719	0.730	0.562	0.566	0.014	0.013
E. Gestational age (weeks)	**-0.033**	**-0.039**	**-0.011**	**-0.015**	**0.008**	**0.004**	**0.017**	**0.015**
Model p-value	0.113	0.063	0.557	0.391	0.623	0.818	0.254	0.293
Resample p-value	0.195	0.139	0.566	0.387	0.610	0.805	0.212	0.271
Romano-Wolf p-value	0.494	0.394	0.785	0.781	0.928	0.936	0.315	0.390
Month-level clustering p-value	0.221	0.144	0.457	0.318	0.269	0.633	0.027	0.043
N	74,796		149,695		228,204		313,251	
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: This table replicates results from Table 3 adjusting for multiple hypotheses testing and using standard errors clustered at conception month level. Multiple hypothesis adjustment uses procedure outlined in Romano and Wolf (2005a) and Romano and Wolf (2005b) and programmed by Clarke et al. (2020). Each estimate, in bold, is based on a separate regression and displays the coefficient on a Reform dummy from Equation 1. We then provide p-value from standard model as in Table 3 followed by resampled p-values, by Romano-Wolf adjusted p-values, and by p-values based on standard errors clustered at month of conception level. When using Clarke et al. (2020) procedure we bootstrap 250 times.

Table A7: Exploring Potential Mechanisms: Heterogeneity Analysis of Maternal Health-related Knowledge and Behaviors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Health-related knowledge					Health-related behaviors during and post pregnancy			
	Small alcohol amounts harmful	Large alcohol amounts harmful	Active smoking harmful	Passive smoking harmful	Knows all of this	Alcohol during pregnancy	Smoking during pregnancy	Smoker postpartum	Exposure to smoking at work
	Panel A. Baseline June 2009 mean (pre-reform period)								
	0.877	0.946	0.969	0.924	0.832	0.129	0.173	0.091	0.114
	Panel B. Interaction with education								
Post 2010-2012	0.128***	0.066***	0.031***	0.086***	0.165***	-0.038***	-0.053***	-0.028**	-0.002
	(0.011)	(0.007)	(0.006)	(0.009)	(0.012)	(0.011)	(0.016)	(0.013)	(0.017)
Post 2010-2012 ×	-0.036***	-0.018***	-0.005	-0.023***	-0.035***	-0.019*	-0.024*	-0.006	0.002
above high school	(0.009)	(0.006)	(0.004)	(0.007)	(0.010)	(0.011)	(0.015)	(0.011)	(0.015)
	Panel C. Interaction with place of residence								
Post 2010-2012	0.113***	0.052***	0.036***	0.074***	0.153***	-0.032***	-0.063***	-0.035***	-0.005
	(0.011)	(0.007)	(0.006)	(0.008)	(0.013)	(0.012)	(0.014)	(0.011)	(0.013)
Post 2010-2012 ×	-0.021	0.002	-0.015**	-0.008	-0.025	-0.036**	-0.014	0.005	0.008
Urban	(0.013)	(0.009)	(0.007)	(0.011)	(0.015)	(0.014)	(0.016)	(0.012)	(0.015)
	Panel D. Interaction with maternal age at conception								
Post 2010-2012	0.135***	0.072***	0.045***	0.105***	0.170***	-0.030	-0.108***	-0.077***	-0.061**
	(0.021)	(0.015)	(0.013)	(0.019)	(0.024)	(0.020)	(0.030)	(0.025)	(0.027)
Post 2010-2012 ×	-0.039*	-0.022	-0.020	-0.040**	-0.035	-0.025	0.043	0.051**	0.068**
Non-teen birth	(0.022)	(0.015)	(0.013)	(0.019)	(0.025)	(0.021)	(0.030)	(0.025)	(0.028)
	Panel E. Interaction with birth order								
Post 2010-2012	0.091***	0.048***	0.024***	0.062***	0.127***	-0.071***	-0.085***	-0.032***	-0.003
	(0.010)	(0.007)	(0.005)	(0.008)	(0.011)	(0.012)	(0.014)	(0.010)	(0.013)
Post 2010-2012 ×	0.020	0.010	0.007	0.014	0.024	0.037***	0.028*	-0.001	0.004
Non-first birth	(0.013)	(0.009)	(0.007)	(0.010)	(0.015)	(0.014)	(0.016)	(0.012)	(0.015)
Observations	9,372	9,372	9,372	9,372	9,372	9,372	9,372	9,372	8,540

Note: Panel A presents mean values of outcome variables in the pre-reform period while panels B to E present regression output where treatment variables are an indicator for survey years 2010 to 2012 and an interaction between this indicator and specific maternal characteristic. In each case we also include the relevant maternal characteristic, however, we do not display these these estimates to conserve space. Each column in panels B to E is based on a separate regression. Additional controls include maternal age at delivery (indicators in three-year intervals), dummies for her marital status, education, province of residence, town size, number of previous births, and a child's gender as well as number of health care providers. Heteroskedasticity robust standard errors in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.