HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 20/19

The Effect of Co-Payments on the Take-Up of Prenatal Tests

Marina Di Giacomo; Massimiliano Piacenza;

Luigi Siciliani and Gilberto Turati

November 2020

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

The Effect of Co-Payments on the Take-Up of Prenatal Tests

Marina DI GIACOMO, Massimiliano PIACENZA,

Luigi SICILIANI, and Gilberto TURATI [*]

November 2020

Abstract

Prenatal screening tests can identify genetic disorders in a fetus, but their take-up remains low in several countries. Using a Regression Discontinuity Design, we test the causal effects of a policy that eliminated co-payments for noninvasive prenatal screening tests in Italy. We find that the policy increases the probability of pregnant women's undergoing prenatal tests by 7 to 10 percentage points, and the effect varies by socioeconomic status. We do not find evidence of substitution effects with invasive tests or that the tests affect the termination of a pregnancy and newborn health. We do find evidence of positive effects on mothers' health behaviors during pregnancy.

Keywords: Prenatal tests, Co-payments, Maternal and newborn health, Regression Discontinuity Design.

JEL Classification: I18, I12, I14.

[*] M. Di Giacomo (Corresponding author): University of Torino, Department of Economics, Social Sciences, Applied Mathematics and Statistics (ESOMAS), School of Management and Economics, Corso Unione Sovietica 218 bis, 10134 Torino, Italy (e-mail: marina.digiacomo@unito.it);

M. Piacenza: University of Piemonte Orientale, Department of Economics and Business (DISEI), Novara, Italy (e-mail: massimiliano.piacenza@uniupo.it);

L. Siciliani: University of York, Department of Economics and Related Studies, York, United Kingdom (e-mail: luigi.siciliani@york.ac.uk);

G. Turati: Università Cattolica del Sacro Cuore, Department of Economics and Finance, Rome, Italy (e-mail: gilberto.turati@unicatt.it).

1 Introduction

Medical literature has long recognized that appropriate prenatal care, such as recommended doctor visits and tests in the first stage of pregnancy, can decrease the likelihood of adverse outcomes for both the mother and the child (e.g., Almond and Currie 2011). Prenatal tests are offered in the first trimester of pregnancy in many high-income countries to identify possible genetic disorders, for example, Down syndrome and Edwards' syndrome, in the fetus (Boyd et al. 2008; Palomaki et al. 2006; Grimes and Schulz 2002). Two different types of test are typically available. Noninvasive *screening* tests, such as combined or integrated prenatal screening tests, are generally inexpensive and allow doctors to estimate the risk of genetic disorders.[1] More invasive *diagnostic* tests are costly but permit doctors to diagnose the presence of chromosomal anomalies with greater accuracy through a genetic map of the fetus.

Combined, these tests can detect congenital anomalies that affect around 0.3 to 0.5 percent of newborns. Congenital anomalies are leading cause of infant mortality, morbidity, and long-term disability (EUROCAT 2010). Accuracy, safety, and the mother's peace of mind are often cited as the benefits of the prenatal tests (e.g., Dormandy et al. 2005). Types of genetic screening programs as well as medical recommendations and guidelines are heterogeneous across countries. At the beginning of the 2000s, many countries implemented new rules to harmonize policies at the national and international levels, but the process is ongoing (World Health Organization 2016).

The literature (e.g., Crombag et al. 2014) and international health organizations (EUROCAT 2010) highlight the variability in take-up rates for prenatal tests across and within countries as a critical issue. Such variability may be explained by many factors, including differences in service delivery and the health system, socioeconomic traits and cultural components of the population, and market failures (e.g., underestimation of benefits and positive externalities associated with the use of prenatal care). Policymakers justify the introduction of free screening programs to the underserved by pointing to the perceived suboptimal take-up of screening tests within some institutional settings (Currie 2006; Shurtz et al. 2016).

In this study, we apply quasi-experimental methods to investigate whether eliminating co-payments for screening tests is an effective policy to increase the take-up of screening and to improve health outcomes for pregnant women and children. We use a regression discontinuity design (RDD) to

[1] Our study does not include the last generation of noninvasive prenatal tests (NIPT), which are based on an analysis of small fragments of DNA of the fetus circulating in the mother's blood. These tests have been offered only in the last five years and exclusively in private facilities.

quantify the effect of a government policy in a large region in Italy that eliminated co-payments for noninvasive prenatal screening tests.

This paper addresses three research questions. First, we test the extent to which eliminating co-payments for noninvasive tests increases their utilization and whether in turn this reduces the utilization of the costlier and riskier invasive tests (a substitution effect). Second, we investigate heterogeneity in the effect across different groups, in particular, women with low socioeconomic status. Although some noninvasive prenatal screening tests are inexpensive, even a relatively small co-payment can be a barrier and discourage women from disadvantaged groups from undergoing the tests. We therefore investigate whether the elimination of the co-payment not only increases utilization but also whether it reduces inequalities in access to prenatal care. Third, the lack of prenatal care is associated with poor birth outcomes, such as low birth weight, preterm birth, and infant mortality (Woodhouse et al. 2014; Currie and Rossin-Slater 2015; Corman et al. 2018), as well as an impaired health status of mothers (Conway and Kutinova, 2006). Thus, we investigate whether an increase in prenatal tests affect newborns' health status and mothers' health behaviors.

We provide evidence on these questions by considering the interesting case of Italy. Regions in Italy have a high degree of autonomy in legislating healthcare policies, including prenatal and maternal care. Prenatal tests for the diagnosis of Down syndrome and other chromosomal disorders have been available since 2001 for all Italian women, including those covered by the Regional Health Service of the Piedmont Region.[2] The regional co-payment scheme required women to contribute to the cost of these tests. The co-payment for noninvasive screening tests was in the range of €27–54, depending on the specific test provided (the combined test, the triple test, or the integrated test). For invasive diagnostic tests, the co-payment was in the €160–200 range, depending on the specific test provided (chorionic villus sampling or amniocentesis) and a number of maternal characteristics.[3]

In 2009, the regional government of Piedmont eliminated the co-payment for noninvasive screening tests, making the noninvasive tests available to all women free of charge. No change occurred to the co-payment for invasive diagnostic prenatal tests. The aim of the policy was twofold. First, public authorities wanted to improve *efficiency* by increasing the take-up of noninvasive screening tests. These tests are less costly and less risky than the more-invasive diagnostic tests. Moreover, recent

[2] Piedmont is a large region located in the northwest of Italy. In 2010, it had around 4.5 million inhabitants, and the number of annual births was 35,000. The Piedmont annual GDP per capita is about €30,000, in purchasing power standards, in 2016 (Eurostat). Piedmont has roughly the same population as Ireland, or a medium-sized US state, such as Kentucky or Louisiana. The crude birth rate (number of live births per 1,000 inhabitants) is around 8, which is slightly lower than the Italian average birth rate (8.5), and much lower than the EU (10), and the US birth rates (12.5).

[3] The Regional Health Service reimbursed hospitals €98 for a noninvasive screening test (e.g., the integrated test) and €592 for an invasive diagnostic test (e.g., amniocentesis).

developments make them almost as reliable as the invasive tests.[4] Screening tests facilitate the identification of women with a high risk of delivering an infant with congenital anomalies. Once identified and informed, these women can obtain the invasive (and costlier for the public purse) diagnostic tests for free. Second, given the observed low take-up rates, mostly among certain more-deprived socioeconomic groups, the policy aimed to improve *equity* by mitigating these disparities in access to prenatal tests.

Our RDD approach to identifying the impact of the co-payment reform is justified by the eligibility rules. First, after the policy became effective, all pregnant women were automatically eligible[5]. Second, the initial general policy was announced in May 2008, but the detailed rules about the policy was made public only in August 2009, which almost coincides with the start of the policy on October 1, 2009. Third, we can reasonably rule out any manipulation of the pregnancy decision as a strategic reaction to the elimination of the co-payment. The decision to have a child is a momentous one, which involves careful consideration of many factors. The price of prenatal tests represents a negligible portion of the cost of raising a child, and the elimination of the co-payment should not have any effect on the number of women becoming pregnant after the policy implementation. This generates a discontinuity in the treatment assignment between women becoming pregnant before and after the policy cutoff date. We therefore use time as our running variable, an approach which is similar to Ito (2015) and Halla et al. (2016).

Our key results are as follows. First, we find that the elimination of the co-payment increases the probability that a woman undergoes any prenatal tests by 7 to 10 percentage points, depending on the functional specification considered. Women's increased use in the screening test, by 7 to 9 percentage points, mostly drives the results. The average pre-policy rate of noninvasive screening prenatal tests was 62 percent; the policy increased the take-up rate for screening tests by 11 to 15 percent. However, we do not observe any substitution effects because the elimination of the co-payment did not reduce the take-up rate of the costlier invasive diagnostic tests. This result is mainly driven by the behavior of women who are older than 35 years at conception. Prior to the elimination of the co-payment, these women could already access free diagnostic tests based on a legislative provision that considered these women to be at high risk of congenital disorders. After the co-payments were eliminated for all women, the over-35 women could also obtain costless screening tests. However, these women have not changed their behavior after the policy change; they are still likely to by-pass the screening tests and undergo the invasive diagnostic tests. Hence, the reduction in the proportion of women who do

[4] Boyd et al. (2008) and Palomaki et al. (2013), for instance, point out that many improvements in prenatal screening methods are responsible for the increase in detection rates of birth defects.

[5] Notice that this makes inappropriate the use of a Difference-in-Differences model: as the policy affects all women, we lack a clear control group.

not undergo any test is due entirely to the increase in the proportion of women who undergo noninvasive screening tests.

Second, we find heterogeneity in the impact of the reform across women with different socioeconomic statuses. The elimination of the co-payment increased the take-up rate, especially for women 25–34 years old, those born abroad, those who reside in metropolitan areas, and those with an intermediate education.

Third, accounting for self-selection in prenatal care, we assess whether screening tests affect the health outcomes of mothers and newborns. After ruling out any increase in the number of pregnancy terminations following the policy, we use the discontinuity in eligibility for free prenatal screening tests triggered by the policy as an instrumental variable for utilization and find that prenatal tests have a positive effect on health behaviors during pregnancy. In particular, screening tests reduce the mother's weight gain, alcohol consumption, and hospital admissions during pregnancy. This is consistent with the view that greater access to prenatal tests may have downstream effects on the use of other prenatal services and on the health outcomes of pregnant women (Metcalfe et al. 2013). In contrast, we do not find any effects of prenatal screening tests on the health status of newborns at delivery.

Prenatal tests have been the focus of economic studies. Fajnzylber et al. (2010) develop and calibrate a model of amniocentesis choice and find that the amniocentesis take-up rate should decrease (instead of increase) with age, once all the risk factors are considered (e.g. the risk of an affected child, the risk of miscarriage associated to amniocentesis, as well as the risk of a decline in fertility with age). Seror (2008) uses an interview-based survey to assess different models of amniocentesis choice. Gajdos et al. (2016) exploit French amniocentesis regulations to offer a measure of the disutility associated with a child with Down syndrome.

The studies by Garrouste et al. (2011) and Shurtz et al. (2016) are the closest to our work. They investigate the effect of public policies aimed at subsidizing diagnostic prenatal tests (amniocentesis) in France and in Israel, respectively. Both studies find that utilization rates rise sharply when women are eligible for full reimbursements.

Our contribution differs from previous studies in three important dimensions. First, our main focus is on noninvasive prenatal screening tests. To our knowledge, no other economics paper has studied this type of prenatal test. Screening tests are not conclusive, but they are less expensive and widely used across different medical specialties (e.g. cholesterol measurement tests, Pap tests, fecal occult blood

tests, mammography) to identify subpopulations that may require additional care.[6] Our study contributes to identifying the effects of the subsidization of prenatal screening tests on take-up rates and, in turn, on mother and newborn health. Second, we test for substitution effects between screening and diagnostic tests, which none of the preceding papers has done. Third, the institutional setting is different, which allows to us to compare findings across health systems.

We also contribute to the literature on prenatal care and its short-run effects on the health of mothers and children. Most of the existing literature on prenatal care focuses on children's health (see Corman et al. 2018 for a recent survey). Only few papers explore, within a causal setting, the beneficial effects of prenatal care on health during pregnancy (e.g. Conway and Kutinova 2006; Yan 2017). Our contribution closes this gap in the existing literature by estimating causal effects of prenatal screening tests on the health of both children and mothers.

The remainder of the paper proceeds as follows. Section 2 presents the institutional background. A conceptual framework to inform the empirical analysis is provided in Section 3. Section 4 describes the data, and Section 5 describes the empirical strategy. In Section 6 we present the results, and Section 7 draws policy implications and conclusions.

2 The Reform of the Co-Payment Scheme for Prenatal Tests

2.1 Institutional setting

The Italian National Health Service (NHS) is a tax-funded system providing universal coverage to all citizens. While funding of the NHS and framework legislation are provided by the central government, the management and provision of day-to-day health services is performed on the regional level (Turati 2013). Within the limits set by national legislation, the twenty regional governments organize their hospital network, staff the public hospitals, and set the Diagnosis Related Group (DRG) rates and the co-payment rules for resident citizens.

In 2000, the Ministry of Health introduced a framework for improving maternal and child health (Ministerial Decree April 24, 2000, *Progetto Obiettivo Materno Infantile* – Project for the Promotion of Maternal, Infant, and Child Health). The plan contained guidelines to define a safe birth plan for mothers and newborns. Recommendations for regional governments included counseling services,

[6] Within the literature on screening programs, Bitler and Carpenter (2016, 2017) find that more-generous insurance coverage induces a large and significant increase in the utilization rates of breast cancer and cervical cancer screening programs (mammography and Pap test, respectively) in the US. Cohen et al. (2015) find that subsidizing a rapid diagnostic malaria test doubled the test take-up rate in Kenya.

prenatal diagnostic care, the de-medicalization of childbirth, active postnatal care, and support for breastfeeding.

One of the recommendations at the national level was to improve the take-up of prenatal genetic testing. Prenatal tests for Down syndrome (trisomy 21) and other genetic diseases (trisomy 18 or Edwards syndrome, and spina bifida) can be grouped into two broad classes: diagnostic and screening tests. Diagnostic tests are more invasive and mainly consist of removing a sample of fluid from the amniotic sac. Given their invasive nature, there is a risk of miscarriage (around one percent). The two most common diagnostic tests are amniocentesis and chorionic villus sampling (CVS), which were the only available prenatal genetic tests for Down syndrome until the mid-1980s. More recently, a number of new prenatal screening tests have been developed. These tests are noninvasive, and they mostly consist of an ultrasound scan (nuchal translucency) and blood tests.[7] The screening tests do not diagnose the presence of a chromosomal disorder, but they estimate the likelihood that the fetus carries certain genetic diseases. These tests are safe for the mother and the fetus, but they lack precision, as false positives or false negatives are possible. If the screening test is positive, the mother must decide whether to seek a definitive (but invasive and risky) diagnostic test.

Most high-income countries offer prenatal genetic screening programs. However, there is significant heterogeneity across countries in policies, recommendations, and guidelines, and in program take-up results (Javaher et al. 2010). Moreover, variability in take-up rates for prenatal tests is explained by differences in service delivery across health systems, as well as socioeconomic traits and culture (Javaher et al. 2010; Palomaki et al. 2013; Vassy et al. 2014; Crombag et al. 2014).

Beginning in the early 2000s, many countries implemented new rules to harmonize national and international policies,[8] but the process is still underway (World Health Organization 2016). The Netherlands and Sweden show the lowest prenatal screening take-up rates (below 30 percent), while Denmark, France, Belgium, and Iceland have the highest (above 80 percent). The United States, the United Kingdom, Finland, and Italy are in between (between 60 and 70 percent), with large regional

[7] The most common noninvasive screening test is the integrated test. It consists of two phases. The first phase (in the first trimester) involves an ultrasound scan (nuchal translucency) and blood tests (the concentration of a chemical called pregnancy-associated plasma protein A, PAPP-A). The second phase involves a blood test where the levels of three markers are measured (alpha-fetoprotein, free beta-hCG, unconjugated estriol). Results from the two phases, along with maternal age, are "integrated" to compute the risk for chromosomal anomalies for that baby. The integrated test will detect eight to nine out of ten pregnancies with chromosomal anomalies, whereas the triple test, consisting of phase 2 of the integrated test only, detects six out of ten. The last generation of noninvasive screening tests (NIPT, also called Non Invasive Prenatal Screening, NIPS or cell free DNA test - cfDNA) have become increasingly available since 2015. These tests are based on an analysis of small fragments of fetus DNA circulating in the mother's blood. In the case of a high probability outcome, a diagnostic invasive follow-up test is still necessary. Our empirical analysis is based on data preceding the introduction of these new screening tests.

[8] Among the others, Boyd et al. (2008), EUROCAT (2010), Javaher et al. (2010), Crombag et al. (2014) and Vassy et al. (2014) survey some recent European national programs. Plachinski (2017) reviews US policies over the last twenty years.

differences *within* countries (Palomaki et al. 2013; Vassy et al. 2014; Crombag 2016). The picture for invasive diagnostic tests is even more fragmented because eligibility varies according to age or risk assessment of genetic disorders. For example, amniocentesis is routinely offered to women older than 35 in the United States, Germany, Sweden, Italy, and Spain, older than 38 in France and Norway, and older than 39 in Finland. In the United Kingdom, only women with a high detection rate, that is, a high risk of genetic disorder based on a screening test, have access to invasive diagnostic tests.

2.2 Policy Reform

Prenatal tests have been routinely offered to all pregnant women in Italy since 2001.[9] Since then, a co-payment, set at the regional level, has been in place, which means that women across Italy must contribute to the cost of these tests. In 2009, the regional government of Piedmont introduced a reform of its co-payment scheme for prenatal screening tests.[10] The new policy made the noninvasive screening tests free for women residing within the administrative borders of the region but maintained the co-payment for diagnostic tests. Prior to the 2009 reform, the co-payment was in the range of €27–€54 for noninvasive screening tests; for invasive tests, the co-payment was €160–€200 depending on the specific test provided (the combined test, the triple test, or the integrated test, and the chorionic villus sampling or the amniocentesis). Some exemptions from the co-payment requirement were available for the invasive diagnostic tests based on the risk of the pregnancy (i.e., if the mother was age 35 or older at the time of conception, there were other cases of chromosomal disorders among members of the family, or the mother had a positive screening test). After the 2009 reform, the co-payment for the noninvasive screening tests was eliminated for all women. Nothing changed in the co-payment scheme for the invasive diagnostic tests: all women had to pay the co-payment, with the only exception being for women with high-risk pregnancies.

Pregnant women can access prenatal tests with a medical prescription from a midwife, a general practitioner (GP), or a gynecologist. The physician must obtain the pregnant woman's informed consent to the genetic screening, and the information provided when obtaining consent must include the characteristics of the test (reliability), the methods of execution, and an exhaustive description of the test. In the case a genetic problem is diagnosed, the physician gives information to the pregnant woman about how best to behave during pregnancy and care for the infant after birth, or about ending the pregnancy. In Italy, the voluntary termination of pregnancy (VTP) is allowed at the woman's request, within gestational limits: it is freely available in the first trimester; it is conditional to save woman's life and protect her physical and/or mental health during the second trimester; it is prohibited

[9] Italian Department for Health, Ministerial Decree, September 10, 1998.
[10] In particular, two resolutions of the Regional Government (Delibere della Giunta Regionale) are relevant: n. 34-8769, May 12, 2008, and n. 38-11960, August 4, 2009.

in the third trimester. As prenatal tests are taken in the first trimester, VTP during the second trimester are those more directly linked to test outcomes. In Italy, according to data provided by the Italian Statistical Office (Istat), only a small share of all VTP occurs in the second trimester, but this share is increasing over time at the national level, probably because of greater use of prenatal diagnosis and the increase in maternal age. The share of second trimester VTP over total VTP was around 2.5% in 2005 and raised to about 5.5% in 2018. Unfortunately, available data do not allow to distinguish about different potential motivations of the VTP: fetal genetic or non-genetic disorders following a test, severe illness of the mother, or other tragic events.

Within our sample, only public hospitals offer prenatal tests because they are the only accredited institutions that satisfy the required international quality standards in the region (AReSS 2008). As only public hospitals can supply the tests, this rules out any monetary incentives for physicians to prescribe prenatal tests. The co-payment flows directly to the hospital, which is also partly paid by the regional government according to a predetermined rate. Physicians working outside the hospitals are paid by capitation, while those working within the hospital are salaried public employees.

A woman's participation in a prenatal program is voluntary and is based on her personal preferences, culture, social background, and medical information (Santalahti et al. 1998). The 2009 co-payment reform did not introduce any new information campaigns or guidelines for obstetricians and gynecologists. Additionally, there was no general advertisement outside the health care medical and administrative staff concerning the co-payment's elimination.

Policymakers justified the reform on two grounds. The first ground was efficiency. Screening tests are robust and efficient in detecting pregnancies at high risk of a genetic disorder. Public health authorities claimed that the take-up rate for these tests was low because the risk assessment was often based only on maternal age.[11] Spurred by recent medical evidence (Wald et al. 2003; Malone et al. 2005; Palomaki et al. 2006; Loane et al. 2013), public authorities aimed to increase the take-up of screening tests, which are characterized by lower costs and lower medical risks than diagnostic tests. The second basis for this change was equity: there was regional evidence of inequalities in the use of prenatal tests based on socioeconomic and cultural traits. This evidence is also supported by the medical literature (Dormandy et al. 2005; Fransen et al. 2010), which reports low take-up rates for some disadvantaged socioeconomic groups. Thus, the aim of the new Piedmont policy was to achieve a higher take-up rate of prenatal screening tests and to mitigate inequalities in access to tests.

[11] In a technical report, the Regional Health Agency (AReSS, 2008) states that a 90 percent take-up rate for prenatal screening tests is a viable, though ambitious, target in Piedmont.

3 Conceptual Framework: Demand for Screening and Diagnostic Tests

To understand the possible effect of the reform on take-up rates, we consider a simple conceptual framework. Define $P^n(c^s, c^d)$ as the proportion of women who do not undergo any test; $P^s(c^s, c^d)$ as the proportion of women who undergo the noninvasive screening test; and $P^d(c^s, c^d)$ as the proportion of women who undergo *only* the invasive diagnostic test, where c^s is the co-payment on the noninvasive screening test, and c^d is the co-payment for the invasive diagnostic test.

Considering standard medical prescriptions, we assume that women who undergo the invasive diagnostic test first never also undergo the noninvasive screening test, since the invasive diagnostic test provides more accurate information than the noninvasive screening one. We also assume that out of $P^s(c^s, c^d)$ who undergo the noninvasive screening test, a proportion Z undergoes only the noninvasive test, while a proportion $(1 - Z)$ also undergoes the invasive diagnostic test. The latter could include women who had a positive test result (conveying "bad news") from the screening test who decide to follow up with the more invasive diagnostic test.

The whole population can therefore be split into four groups, which could be interpreted as related to the demand for each of the tests, or for both tests:

- $P^n(c^s, c^d)$ do not undergo any test,
- $P^d(c^s, c^d)$ undergo *only* the invasive diagnostic test,
- $Z \times P^s(c^s, c^d)$ undergo *only* the noninvasive screening test,
- $(1 - Z) \times P^s(c^s, c^d)$ undergo *both* tests,

where $P^n + P^d + P^s Z + P^s(1 - Z) = 1$, so that $P^n = 1 - P^d - P^s$.

Assume that $\partial P^s/\partial c^s < 0$, $\partial P^s/\partial c^d \geq 0$: an increase in the co-payment for the noninvasive screening test reduces the proportion of women who undergo the noninvasive screening test; and an increase in the co-payment for the invasive diagnostic test increases the proportion of women undergoing the noninvasive screening test if the tests are substitutes.

Similarly, $\partial P^d/\partial c^d < 0$, $\partial P^d/\partial c^s \geq 0$: an increase in the co-payment for the invasive diagnostic test reduces the proportion of women who undergo only the invasive test; and an increase in the co-payment for the noninvasive screening test increases the proportion of women who undergo the invasive diagnostic test if the tests are substitutes.

How does the elimination of the co-payment for the noninvasive screening test affect the four groups identified above?

1) $\partial(ZP^s)/\partial c^s < 0$. The elimination of the co-payment for the noninvasive screening test *increases* the proportion of women who undergo *only* the noninvasive screening test.

2) $\partial P^d/\partial c^s \geq 0$. The elimination of the co-payment for the noninvasive screening test weakly *reduces* the proportion of women who undergo *only* the invasive diagnostic test.
3) $\partial((1-Z)P^s)/\partial c^s \leq 0$. The elimination of the co-payment for the noninvasive screening test weakly *increases* the proportion of women who undergo *both* tests.
4) $\partial P^n/\partial c^s = \partial(1-P^s-P^d)/\partial c^s = -\partial P^s/\partial c^s - \partial P^d/\partial c^s$. The elimination of the co-payment for the noninvasive screening test weakly *decreases* the proportion of women who do not undergo any test if the increase in the proportion of women who undergo the noninvasive screening test is higher than the decrease in the proportion of women who undergo the invasive diagnostic test.

Finally, note that the elimination of the co-payment for the noninvasive screening tests always increases the proportion of the population undergoing the noninvasive screening test, P^s, that is, the sum of those who undergo only the noninvasive test or both tests. The effect on the proportion undergoing the invasive diagnostic test, $(P^d + (1-Z)P^s)$, that is, the sum of those who undergo only the invasive test or both, is instead indeterminate and given by $\partial P^d/\partial c^s + (1-Z)\,\partial P^s/\partial c^s \gtrless 0$. On one hand, the elimination of the co-payment for the noninvasive screening test encourages a reduction in the proportion of those who undergo only the invasive diagnostic test due to the substitution effect, (the first term above), but also increases the proportion of those who undergo both tests (the second term above), due to a higher proportion of positive test results (conveying "bad news"), which in turn induces more women to undergo the invasive test. These results therefore highlight the usefulness of splitting the whole population into four groups (i.e., no test, only noninvasive screening test, only invasive diagnostic test, both tests), rather than three groups (i.e., no test, noninvasive test, and invasive test). In the empirical section, we test these four predictions.

4 Data

4.1 Sources and sample definition

We exploit the CEDAP (CErtificato Di Assistenza al Parto, literally "Delivery Certificate") administrative archive. The Delivery Certificate was introduced by the Ministry of Health in 2001 (Ministerial Decree n. 249, July 16, 2001). It allows officials to collect homogenous, comparable, and high-quality statistical data related to all births occurring in public and private hospitals or any other facility (e.g., midwifery units, birth centers, home births) in Italy. The midwife or the doctor who attended the birth is responsible for filing the form within ten days of the delivery. The compulsory certificate contains a rich set of variables about sociodemographic characteristics of the parent(s), the course of the pregnancy, labor, childbirth, and the health status of the newborn. The certificate also contains data on prenatal testing, in particular whether the mother underwent a prenatal test and which

type of prenatal tests. The certificate does not report the exact date the woman undergoes the prenatal test, neither the result of the test. The sources of all data in the certificate are medical records and official personal data, except for the socio-economic information (marital status, education level, and employment status) that are self-reported.

Our data include women who actually give birth.[12] The data cover around 98 percent of all births within the Region. Home births (representing 0.2 percent of total births), and deliveries in private hospitals (1.8 percent of total births) are excluded.

We define our final sample according to the following criteria. First, we consider all women who became pregnant one year before or one year after the elimination of the co-payment, which occurred on October 1, 2009. They all delivered in one of the 32 public hospitals within the administrative area of Piedmont. We are primarily interested in the conception dates, which of course, is an estimate. In our sample, the calculation of the conception date is based on the gestational age of the fetus, which, in turn, is based on several fetal measures obtained at the first ultrasound exam. Ultrasound exams have an intrinsic error of up to five to seven days (Verloove et al. 1986). For this reason, we choose "week" as the observational unit for our forcing variable (i.e., the number of weeks from conception to/from the elimination of the co-payment). We then exclude women whose date of conception falls in the twelve weeks between the beginning of July 2009 and the end of September 2009. Since we do not observe the exact date when the women undergo the prenatal test, it may be that some of the women whose conception date falls in the period of July 2009 through September 2009 actually paid the co-payment, while some others did not pay it. Medical guidelines establish that screening tests are delivered between the 11th and the 13th gestational week, and we safely exclude pregnancies that are borderline with respect to the co-payment elimination policy.[13]

The total number of observations at this step is 71,283 women conceiving 52 weeks before 1st July 2009 and 52 weeks after the policy change, i.e. 1st October 2009. Then, we restrict our sample to nulliparous women who do not exhibit any pathological disorders during pregnancy (high blood pressure, diabetes, placenta praevia, psychological disease, familial occurrence of congenital malformation, etc.). Following previous literature, the exclusion of women who experienced past pregnancies (we excluded 32,950 observations here) allows us to avoid some confounding effects of

[12] We do not have information on women terminating their pregnancy before delivery occurs. Pregnancies mainly end before delivery because of spontaneous abortions, or because of induced (or voluntary) abortions. In Piedmont, the former represents around 11 per cent of all pregnancies, while the latter about 15 percent of all pregnancies. Information is from Regione Piemonte, Assessorato Sanità.

[13] The exclusion of observations immediately surrounding the threshold is also known as "donut RDD" (see Barreca et al 2011). However, our results are robust to the inclusion of all women; when all women are included, those who become pregnant between July and the end September 2009 are assumed to fall in the pre-policy period. Results are also robust to the exclusion of women whose conception dates fall within different time windows (e.g. from eight to sixteen weeks before the elimination of the co-payment).

the policy, as past experience may have a prominent role in the choice of prenatal tests (and prenatal care behavior, in general). The reason for excluding women with any pathological disorders (we exclude 5,504 observations) is to remove all women who, in the pre-2009 period, had access to free prenatal tests due to their high-risk pregnancies.[14]

After applying our inclusion criteria, the final sample consists of 32,829 women: 52 percent of them became pregnant during the pre-reform period (17,053 women), while the remaining 48 percent became pregnant in the post-reform period (15,776 women).

4.2 Variables definitions and summary statistics

Our first set of dependent variables includes the prenatal tests take-up rates. Consistent with our conceptual framework, we consider the four following utilization rates:

- No test (P^n): a binary indicator equal to one if the woman had no prenatal tests during her pregnancy, and zero otherwise;
- Screening (noninvasive) test only (ZP^s): a binary indicator equal to one if the woman had only a noninvasive prenatal screening test, and zero otherwise;
- Diagnostic (invasive) test only (P^d): a binary indicator equal to one if the woman had only an invasive diagnostic prenatal test, and zero otherwise;
- Both tests ($(1-Z)P^s$): a binary indicator equal to one if the woman had a prenatal screening test followed by a diagnostic test, and zero otherwise.

Table 1 shows summary statistics for the take-up rate of prenatal tests in the whole sample and in the pre- and post-policy periods. By comparing the pre-reform period to the post-reform period, the percentage of women taking up noninvasive screening tests increases by around 9 percentage points (from 62 percent in the pre-reform period to 71 percent in the post-reform one), while the rate of invasive diagnostic tests raises by less than 1 percentage point (from 6.7 percent to 7.3 percent). The percentage undergoing no tests decreases by 10 percentage points: from 28 percent to 18 percent after the policy change, while the take-up rate for those undergoing both tests only marginally increases (0.2 percentage points).

The second set of dependent variables includes women's and newborns' health status.

For the health status of mothers, we exploit information about lifestyle during pregnancy: weight gain during pregnancy, smoking, alcohol consumption, the use of folic acid, and hospital admissions during pregnancy. For the health status of the newborns, we use the following indicators, all measured

[14] We also experiment with including these women (multiparous women and/or women with pathological conditions) in our estimated models, and the results do not qualitatively change.

at birth: newborn weight, length, and head circumference at birth, the Apgar scores after one minute and after five minutes from birth, a resuscitation binary indicator, preterm birth, and stillbirth. Table 2 reports the summary statistics for the health outcomes of women (Panel A) and of newborns (Panel B).

International guidelines provide guidance on the optimal weight gain during pregnancy (Institute of Medicine 2009). For a woman of normal weight, weight gain during pregnancy should not exceed 11–13 kg (24–29 pounds). Excessive weight gain is associated with poor health outcomes such as hypertension, gestational diabetes, miscarriage, and delivery complications. In our analysis, we use two measures of weight gain: a continuous variable (Weight Gain in Pregnancy) expressed in kg, and a dichotomous variable of excessive weight gain (Weight Gain in Pregnancy>15 kg) equal to one if the weight gain during pregnancy is higher than 15 kg (see Currie et al. 2010). The average weight gain during pregnancy in our sample is 13.1 kg, with a lower average in the post-policy period (13 kg) relative to the pre-policy year (13.2 kg). Around one quarter of women gain more than 15 kg; this proportion is slightly smaller in the post-policy period (23.5 percent versus 24.5 percent in the pre-policy period).

Smoking and consuming alcohol during pregnancy are both associated with health problems for the fetus, including growth problems and damage to the nervous system. In our sample we exploit two dummy variables for smoking and alcohol consumption. The variable “Smoke in Pregnancy” is equal to one if the woman declares during pregnancy that she smokes at least one cigarette per day, while the variable “Alcohol in Pregnancy” is equal to one if the woman acknowledges consuming at least 12 g of alcohol per day (about one glass of wine per day) during her pregnancy. Around 7 percent of the sample smoke during pregnancy (7.1 percent in the pre-policy period and 7.5 percent in the post-policy), and 4 percent consume alcohol (5.4 percent in the pre-policy period and 2.8 percent in the post-policy period).

A deficiency of folic acid (a type of B vitamin) during pregnancy is related to fetal malformation, and folic acid supplements are recommended for all pregnant women and for women who plan to become pregnant. We consider a binary variable “Folic Acid in Pregnancy” equal to one if the prospective mother used folic acid during her pregnancy: 83 percent of women, on average, use folic acid during pregnancy, with a larger share in the post-policy period (86 percent versus 79 percent in the pre-policy period).

A dichotomous variable for hospital admissions during pregnancy is defined as equal to one if the woman experiences one or more hospital admissions during pregnancy. Around 4 percent of women

experience hospitalization at least once during the pregnancy. The average share is larger in the post-policy period (3.5 percent pre-policy versus 3.9 percent post-policy).

Birth weight is a widely accepted measure of newborn well-being and a low birth weight is associated with a higher probability of mortality and morbidity, both in the short and long runs (e.g. Almond et al. 2005). We follow the World Health Organization's definition of "Low weight" as a binary variable, equal to one if the birth weight is below kg 2.5 (roughly 5.5 pounds). On average 5.5 percent of newborns have low birth weight. Similarly, the newborns' length and head circumference are related to their well-being. The average length is around 49.5 centimeters (19.5 inches), while the average head circumference is 34 centimeters (13.4 inches).

The Apgar score is a common measure of the physical condition of a newborn. The Apgar score, when measured at one minute after birth, assess how well the infant tolerated the birthing process, while the Apgar score measured at five minutes after birth assesses how well the infant is adapting to life outside the womb. The score ranges from 0 (no vitality) to 10 (high vitality) and is measured by doctors and nurses by points for heart rate, respiratory effort, muscle tone, response to stimulation, and skin coloration. Apgar scores have been found to be predictive of health status, cognitive ability, and children's behavioral problems (Almond et al. 2005; Figlio et al. 2014). In our sample we define two "Low Apgar Score" indicators, at one and at five minutes, as binary variables equal to one if the Apgar score is below 9. In our sample, 20 percent and 5.5 percent of newborns have low Apgar scores after one minute and after five minutes, respectively.

The resuscitation binary indicator is equal to one if a resuscitation method (drugs, intubation, cardiac massage, or oxygen at birth) was necessary after birth. In our sample, a resuscitation method was used in 3.4 percent of deliveries.

The preterm binary indicator is equal to one if the birth occurred before the 37^{th} gestational week. A premature birth is a strong predictor of mortality and morbidity and is associated with poor health (Borra et al. 2016). We observe preterm birth in 5.4 percent of newborns.

Finally, stillbirth (death of the infant immediately before or during delivery) occurs in 0.3 percent of deliveries.

In all specifications we include the following information about the woman: age, education level, employment status, the employment status of her partner, marital status, whether she had a twin pregnancy, previous miscarriages and abortions, residence, and nationality.

Table A1 in the Appendix provides a detailed description of all of the women's characteristics we consider, while Table A2 shows some summary statistics. Around one-third of women are between

ages 30–34 years old when they conceive, 17 percent are younger than age 25, 28 percent are ages 25– 29, and the remaining 21 percent are older than 35. While 50 percent of women have a "medium" level of education (high school), 28 percent of the sample have a "low" level of education (compulsory education or less), and the remaining 22 percent have a "high" level of education (tertiary education or more). In the post-policy period we observe a higher rate of women with high education (23 percent in the post-policy period versus 20 percent in the pre-policy period) and a lower percentage of women with low education (27 percent in the post-policy period versus 29 percent in the pre-policy period). Around 70 percent of women are employed, and 60 percent are married. More than three-quarters of women are Italian nationals. A working father is observed in 90 percent of cases, while twin pregnancies occur in 1 percent of the sampled women. The proportion of women who experienced past miscarriages and abortions is about 12 percent and 8 percent, respectively. Women living in metropolitan areas[15] represent around 40 percent of the sample. The proportion living in metropolitan areas slightly increases in the post-policy period (41 percent post-policy versus 37 percent pre-policy).

Concern about inequalities in access to prenatal tests was one of the reasons policymakers gave for eliminating the co-payment. Table 3 shows descriptive evidence about the proportion of women who did not undergo any prenatal tests across several demographic and socioeconomic traits of mothers.

For all age groups, the proportion of women who do not undergo any test decreases by about 10 percentage points after the policy change. The proportion undergoing prenatal tests increases with age; this is consistent with medical guidelines promoting these tests for older mothers, given their increased risk of giving birth to babies with congenital defects.

The proportion of women with low education who do not undergo any test decreases by about 11 percentage points in the post-policy period, while for women with medium and high education, the proportion decreases by 9 percentage points and 8 percentage points, respectively.

Native and foreign women react to the policy change in a similar way, and the proportion not undergoing any test decreases by 10 percentage points for both groups. Finally, the reduction is larger for women living in metropolitan areas relative to nonmetropolitan areas.

[15] Metropolitan areas are defined by the metropolitan area of Torino—the regional capital—and of the other seven largest towns of the region (provincial capitals). Each metropolitan area includes the municipality of a medium-large town and the neighboring municipalities, which depend on the main town for most services (healthcare, education, etc.). This may cause difficulties in accessing the services due to user congestion and possible deficiencies in the connections between the center and the periphery (e.g., poor public transit systems), especially for those citizens living in the suburbs.

5 Estimation Strategy

Our estimation strategy is twofold. First, we test the effect of eliminating the co-payment on the take-up of prenatal tests and the extent to which the effect differs by socioeconomic status. Second, we investigate whether the variation in the utilization of prenatal tests affects mothers' health behaviors during pregnancy (weight gain, smoking, alcohol consumption, folic acid supplements, hospitalization) and newborns' health outcomes (birth weight, body length, head circumference, Apgar scores, resuscitation, preterm birth, and stillbirth).

5.1 Prenatal tests take-up rates

To analyze the effect of eliminating the co-payment on prenatal tests utilization, we adopt an RDD approach. We use time, measured in weeks, as the forcing variable (e.g. Ito 2015; Halla et al. 2016; Hausman and Rapson 2018). In this study, the time threshold determines eligibility for the elimination of the co-payment for prenatal screening tests. The eligibility rule has three key features. First, women who become pregnant after the threshold date (the policy change) are automatically eligible, while those becoming pregnant before the threshold date are ineligible.[16] Second, the co-payment was eliminated shortly before the cutoff date: a general announcement was made in May 2008, but the detailed rules were not set forth until August 2009, which almost coincides with the start of the new policy (October 1, 2009). Third, the price of prenatal tests (€27–€54) represents a negligible portion of the cost of raising a child. We can reasonably rule out that women strategically delay pregnancy to avoid the co-payment. Having a child is a momentous decision that involves careful consideration of many factors. It follows that the pregnancy decision is exogenous with respect to the threshold date that generates a discontinuity in the treatment assignment.

Our sample includes women at the date of conception. If the conception date falls in the post-policy period, the woman is automatically enrolled in the program and she can obtain a screening test for free. The time threshold simply determines which eligibility rule was in effect. Since we plausibly exclude strategic behavior (women do not delay/anticipate pregnancy as a consequence of the co-payment elimination), our identification is based on asymptotics in N, the number of cross sectional units around the threshold, rather than asymptotics in T, the number of data points in time. This allows us to exploit the standard cross-sectional RD toolkit (Ito 2015; Hausman and Rapson 2018).

[16] We could perform our analysis using an interrupted time series (ITS) (Kontopantelis et al. 2015) or an event-study. We are dealing with a health policy targeted at the population level, so we lack a clean control group. Hence, in principle, both ITS and the event-study approach would not be ill-suited. However, note that unlike an ITS, our main unit of analysis is the individual woman and not a "time frequency" (e.g., daily observations). The main identification assumption is that both before and after the policy change, women only differ in eligibility.

There are two main strategies for specifying the functional form to estimate the magnitude of the discontinuity in the outcome at the cutoff point within a classical RD setting: the parametric approach and the nonparametric approach (Imbens and Lemieux 2008; Lee and Lemieux 2010). The main difference is in the way data around the cutoff are used: while the parametric approach focuses on the optimal functional form to fit the full set of data, most nonparametric approaches search for the optimal data bandwidth where a linear regression function can produce a consistent estimate.

We mainly apply a parametric approach and we fit flexible parametric functions to data within a one-year interval around the cutoff point (conception dates 52 weeks before and 52 weeks after the policy implementation). As a robustness check, we also implement a number of local linear and polynomial (up to the fourth order) nonparametric specifications (Calonico et al. 2014, 2015, 2018) where we use a range of different methods to choose the "optimal" bandwidth (Imbens and Lemieux 2008; Lee and Lemieux 2010; Calonico et al. 2015).

The parametric specification is as follows:

$$(1) \qquad Y_i = \alpha_0 + \delta A_i + \sum_{k=1}^{K} \alpha_k (C_i - C^*)^k + \sum_{k=1}^{K} \lambda_k A_i \times (C_i - C^*)^k + X_i\theta + \varepsilon_i$$

where K can take the following values $K = 1, 2, 3, 4$. We define the dependent variable Y_i as a binary indicator identifying in turn each of the four possible utilization rates classified in our conceptual framework in Section 3. First, the dependent variable Y_i is equal to one if woman i undergoes no test (P^n), and zero otherwise. Then, we re-estimate equation (1) by using as dependent variable Y_i , which is equal to one if woman i undergoes only the noninvasive subsidized screening test (ZP^s), and zero otherwise. In order to assess the presence of substitution effects, we also estimate equation (1) after replacing Y_i with a variable equal to one if woman i undergoes only the invasive diagnostic test (P^d), and zero otherwise. Finally, Y_i equals one if woman i undergoes both types of test ($(1 - Z)P^s$), and zero otherwise.

C_i is the (presumed) conception date for woman i, while C^* is the cutoff date. The difference $(C_i - C^*)$ measures the number of weeks between the conception date and the cutoff date. A_i is the treatment assignment dummy variable, the primary variable of interest: it takes a value of one in the post-policy period ($\mathrm{A_i} = 1$ if $\mathrm{C_i} \geq \mathrm{C^*}$), and zero otherwise. The coefficient δ is our key coefficient of interest and gives us the change in the likelihood of undergoing the prenatal test at $C_i = C^*$.

X_i is a vector of observable characteristics of woman i such as age, highest education level, parents' employment status, marital status, twin pregnancy, previous miscarriages and abortions, area of residence, and nationality. ε_i is the error term.

Throughout, we cluster the standard errors at the level of the mother's district of residence–quarter of conception, to account for any correlation within the clusters.[17] The Piedmont territory is divided into 14 Local Health Authorities (LHA, Aziende Sanitarie Locali), and each authority is responsible for providing health services in its assigned geographical area or district. Moreover, we observe mothers over ten quarters. By interacting the two variables (district of residence or LHA of the woman and quarter of conception), we obtain 138 clusters of data.[18]

Identification of the model requires no self-selection at the cutoff and no discontinuous differences in the characteristics of women at the cutoff date. We discuss both issues in Section 6.1 below. To address heterogeneity issues, we also estimate equation (1) for a set of subsamples of mothers according to their demographic and socioeconomic characteristics (age, level of education, nationality, and area of residence).

5.2 Women's and newborns' health outcomes

Prenatal care may affect the health of both the mother-to-be and the newborn. Prenatal care may affect the health of mothers directly, for example, through hypertension, hospitalizations, or mental health (Yan 2017), but also indirectly, through a change in lifestyle (Conway and Kutinova 2006) or the decision to voluntarily terminate pregnancy (Natoli et al., 2012, Jacobs et al., 2016). Lack of prenatal care is also associated with poor birth outcomes, such as low birth weight, preterm birth and infant mortality (Woodhouse et al. 2014; Currie and Rossin-Slater 2015; Corman et al. 2018).

We analyze whether the take-up of prenatal tests, which is a measure of prenatal care use,[19] has any consequences on health behaviors during pregnancy or on newborns' health. Specifically, we investigate whether improved access to prenatal screening tests has downstream effects on the health outcomes of women and newborns. These effects might be explained by the behavioral channel of the new policy improving prenatal care (e.g. through a greater use of related prenatal care services, Metcalfe et al. 2013), which in turn affects health outcomes.

Our dataset includes information related to lifestyle during pregnancy: weight gain, smoking, consuming alcohol, using folic acid, and hospitalization[20]. We measure the general health of

[17] In terms of standard errors, the results are qualitatively similar when we cluster by week of conception, that is, the measurement unit of the conception date C_i, as suggested by Lee and Card (2008) for discrete running variables.
[18] The data span from the first week of July 2008 to the first week of October 2010 (with the exclusion of the data from July through September 2009, see footnote 12), that is, the ten quarters from the third quarter of 2008 to the fourth quarter of 2010. There are no observations in two clusters.
[19] Other prenatal care measures include the number of prenatal care visits, the delay in initiating prenatal care, prenatal education, and the number and type of prenatal tests and scans (Currie and Rossin-Slater 2015; Corman et al. 2018). Unfortunately, our dataset does not contain any of these measures.
[20] Since we do not have information on pregnancies ending without a delivery, we are unable to provide an analysis of the impact of the policy of VTP. However, we discuss below aggregate regional data on VTP during the second trimester from which no effect can be found when the reform kicks in.

newborns at delivery by birth weight, length, head circumference, Apgar scores, whether the newborn was resuscitated, premature birth, and stillbirth.

We estimate the following regression model on the sample of women and newborns:

$$HS_i = \beta_0 + \beta_1 Y_i + \beta_2 (C_i - C^*) + \beta_3 A_i \times (C_i - C^*) + X_i \eta + u_i \quad (2)$$

where the health status HS_i of the infant/mother *i* depends on Y_i, a dummy variable equal to one if the mother had prenatal screening tests, and zero otherwise. In the specification we also include the distance to the cutoff point $(C_i - C^*)$, the treatment assignment dummy A_i interacted with the distance to the cutoff point $(C_i - C^*)$, and the full set of mother characteristics X_i, from equation (1).

We analyze this relationship within a fuzzy RD design (Lee and Lemieux 2010). Utilization of prenatal tests is voluntary, and it is usually based on a woman's preferences and sociocultural background. Self-selection into prenatal care may bias our results upward if more-diligent and healthier women are more likely to use prenatal care, or downward if women expecting worse health demand more prenatal care. We tackle this issue using a 2SLS strategy where the instrumental variable for prenatal test utilization is the discontinuity in the eligibility of free prenatal screening tests triggered by the new policy. The policy that eliminated the co-payment affects the probability of undergoing a prenatal test, but it does not directly affect health outcomes and therefore offers exogenous variation that we exploit in 2SLS estimation. Equation (2) is thus estimated via 2SLS, where the first stage is:

$$Y_i = \alpha_0 + \delta A_i + \alpha_1 (C_i - C^*) + \lambda A_i \times (C_i - C^*) + X_i \theta + \varepsilon_i,$$

which is equation (1) with $K = 1$, that is, the linear specification. The coefficient of interest is β_1 in equation (2), which gives a measure of the causal effects of prenatal test utilization on women' health-related behaviors and newborns' health. More specifically, we measure the average causal effect of screening tests on compliers, that is, the subpopulation of women randomly assigned to treatment (the co-payment elimination group) who comply with the assignment by undergoing a screening test.[21] Finally, we investigate heterogeneous effects by estimating equation (2) for subsamples of mothers based on age, education, nationality, and area of residence.

[21] Our setting is characterized by partial compliance. The population can be divided into three subgroups: (i) compliers: women randomly assigned to treatment, who comply with it by undergoing a prenatal test; (ii) always-takers: women who are not assigned to treatment, who do undergo the prenatal test; (iii) never-takers: women randomly assigned to treatment, who do not undergo a prenatal test. Our IV strategy allows us to identify the effect of prenatal tests on the group of compliers (see Angrist and Pischke 2009).

6 Estimation Results

6.1 Statistical tests and descriptive evidence

We present two sets of statistical tests: sorting of women around the cutoff and balancing tests for the comparability of women around the cutoff. First, we test for strategic manipulation of the conception date: if women nonrandomly sort themselves around the cutoff date, for example, by delaying the conception date to obtain free access to the noninvasive screening test, the continuity assumption of average potential outcomes does not hold and the causal effect is not identified. We perform several statistical tests for the presence of any discontinuities in the density around the cutoff date based on local-polynomial density estimation techniques (McCrary 2008; Cattaneo et al. 2018). We do not find any statistically significant evidence of manipulation. We find no discontinuities in the density of women conceiving around the cutoff date, which supports the absence of self-selection or nonrandom sorting of women into control and treatment groups.[22]

Next, we test for changes in the observable characteristics of women around the cutoff date, and therefore check the smoothness of the control variables around the policy change date. Table A3 in the Appendix reports the estimated coefficients for the treatment assignment dummy only. None of the characteristics of mothers show any statistically significant discrete jump at the cutoff date (except for past miscarriages). Figure A2 in the Appendix plots the observed characteristics of pregnant women against the conception week, relative to the eligibility cutoff date. Also from graphical inspection, we find that observable mothers' characteristics are smoothly distributed around the cutoff date.

Finally, we present graphical evidence for our outcomes. Figure 1 plots the four prenatal tests utilization rates of the four prenatal tests against the assignment variable, that is, number of weeks separating the conception date from the cutoff date. The y-axis measures the proportion of women undergoing no prenatal tests (top-left panel), a screening test only (top-right panel), a diagnostic test only (bottom-left panel), and both a screening and a diagnostic test (bottom-right panel). The x-axis reports $(C_i - C^*)$, that is, the number of weeks between the conception date for woman i C_i , and the cutoff date C^*. The zero value represents the cutoff date of the policy change. All women whose conception date falls before the cutoff, do not access free prenatal screening tests, while all women whose conception date falls after the cutoff can access for free the noninvasive screening prenatal test.

[22] Figure A1 in the Appendix shows the density distribution (histogram) of women around the cutoff date.

If the elimination of the co-payment had an effect, we expect a discontinuous jump in the utilization of prenatal tests at the cutoff date. Figure 1 documents a clear discontinuity in the probability of undergoing prenatal tests as a function of the conception date in two out of four groups. In particular, in the first two top panels of Figure 1, there is a discontinuous change in the utilization rates at the cutoff point: we observe a reduction in the probability of undergoing no prenatal tests after the program implementation and an increase in the take-up rate for noninvasive screening tests. For the proportion of diagnostic tests and of both tests (two bottom panels of Figure 1) we do not observe any significant change at the cutoff date.

6.2 Effect on prenatal tests take-up

Table 4 presents the RDD estimates for the effect of the elimination of the co-payment on the probability of undergoing prenatal tests. Results are based on the estimation of equation (1) by OLS on the whole sample of nulliparous women who become pregnant within a time bandwidth of 52 weeks around the policy intervention. Standard errors are clustered at the level of the district of residence (LHA)–quarter of conception of the mother (138 clusters); the specification is linear. We first estimate the effect of the policy change on the decision to undergo a prenatal test or not (outcome "No test," in column 1). We then estimate separately the effect on the probability of undergoing a screening test (column 2), a diagnostic test (column 3), and both screening and diagnostic tests (column 4).

The results show that the policy had a positive statistically significant effect on the utilization rate, and this effect is mainly due to screening test take-up rates. The coefficient for the treatment assignment variable A_i is negative and statistically significant in column 1: the probability of undergoing no prenatal tests decreases by around 7 percentage points after the elimination of the co-payment. We also control for observable characteristics in all specifications. Women who are older, more educated, employed, and native show a higher probability of undergoing prenatal tests. Similarly, we find a significantly higher probability of undergoing tests for women who experienced past miscarriages and abortions. Married women and women living in metropolitan areas have a higher probability of undergoing no prenatal tests.

In column 2 of Table 4, we find that after the elimination of the co-payment, the probability of undergoing a noninvasive screening test increases by 6.5 percentage points. The probability of having an invasive diagnostic test only, and the probability of undergoing a screening test followed by a diagnostic test, are not affected by the policy.

We extend our baseline results on several dimensions as robustness checks. First, we introduce higher-order polynomial parametric specifications. Table A4 in the Appendix shows OLS estimation

results for equation (1) with different nonlinear trends (quadratic, cubic, and quartic in columns 1, 2, and 3, respectively). The policy triggered an increase in the take-up of prenatal tests (by 9 to 11 percentage points) that is due to the increase in the probability of undergoing screening tests by around 9 percentage points. Diagnostic tests are not affected by the treatment assignment.

Second, we assess the robustness to different bandwidth choices around the cutoff date. Table A5 in the Appendix shows that the main results are unchanged, with both halving the bandwidth to 26 weeks and doubling the bandwidth to 104 weeks around the cutoff date.

Third, we estimate several local linear and polynomial (up to the fourth order) nonparametric specifications. Table A6 in the Appendix reports results for the nonparametric estimation method proposed by Calonico et al. (2014), where we allow for the selection of the optimal bandwidth by the MSE-optimal bandwidth selector, and the standard errors are still clustered robustly at the level of district of residence (LHA)–quarter of conception. Coefficients are similar in size to the OLS results in Table 4 and Table A4. Reassuringly, the results are robust to polynomial specifications of higher-order (third-order and quartic-order) polynomials. We find that after the policy change, the probability of undergoing no test significantly decreases by around 9 to 10 percentage points, while the take-up rate for screening tests significantly increases by 8 percentage points. There is no significant change for women who undergo only the diagnostic test or both tests.

Finally, Figure A3 in the Appendix presents a falsification test. We artificially move cutoff date backwards to verify that the treatment assignment effect is zero at some placebo thresholds. We do not find any evidence of significant discontinuity before the policy was actually implemented.

Overall, our evidence points to an increase in the take-up rate of screening tests following the elimination of the co-payment: the probability that a woman undergoes any prenatal tests increases by 7 to 10 percentage points depending on the considered specification. This result is driven by the increase in the take-up of screening tests (by 7 to 9 percentage points). Evaluated at the average pre-policy rate of 62 percent, the policy increased the take-up rate for screening tests by 11 to 15 percent. However, as the elimination of the co-payment did not affect the utilization rate for the riskier and costlier diagnostic tests, we do not find evidence of any substitution effects.

6.3 Heterogeneity and long-run effects

In this section we explore whether the policy had more pronounced effects on prenatal test take-up among specific subgroups who differ by age, education, nationality, and residence. Then, we assess whether the effects of the policy persist over time.

We split the sample of mothers into four age groups based on the date of conception: 18–24, 25–29, 30–34, or over 35 years. Women who are 35 years old or older at conception have free access to invasive diagnostic prenatal tests, because they are considered at high risk of congenital disorders. The elimination of the co-payment for this age group was introduced at the national level in 2001, which was well before the period we study (2008–2010). After the elimination of the co-payment for screening tests, the over-35 women had free access to any screening or diagnostic prenatal tests. We are particularly interested in testing whether this group of older women changed any testing behavior after the policy change given the higher possibility of a substitution effect between the invasive and the noninvasive tests as they both became fully subsidized.

Table 5 shows the results for the OLS estimation of equation (1) on different subsamples of women according to their age at conception. Standard errors are clustered at the level of district of residence–quarter of conception. The odd-numbered columns show the linear specifications, and the even numbered columns the quadratic ones.

In columns 1 to 2 of Table 5, we find that the probability that a woman does not undergo any test decreases by 7 to 10 percentage points if she is below age 35 at conception. Columns 3 and 4 show that this result is coupled with a higher probability of screening tests, which increases by 7 to 11 percentage points. No statistically significant change occurred to the probability of undergoing diagnostic tests or both tests. For the 18–24 age group, the sign and the magnitude of the coefficients are similar to the 25–29 and 30–34 age groups, but the standard errors are larger, possibly due to smaller sample sizes.

We find that for women who are 35 or older at conception, the probability of undergoing any prenatal test increases by 8 to 9 percentage points (columns 1 and 2 of Table 5). However, unlike other age groups, this effect is not completely determined by the increase in the noninvasive screening tests. Both the take-up rates of screening tests (columns 3 and 4 of Table 5) and of diagnostic tests (columns 5 and 6 of Table 5) increase after the policy implementation, even if neither of the effects is statistically significant. A possible rationalization of this result is a behavioral response to the subsidized screening test by women who also have free access to subsidized diagnostic tests (Cohen et al. 2015). After the elimination of the co-payment, women may be more prone to collect additional information about prenatal tests in general. Women with higher risks, for whom all prenatal tests are free, end up choosing the conclusive diagnostic invasive test more frequently.

To further explore the effect of the policy on the testing behavior of women over 35, we consider a smaller subsample defined by those women who are 32–38 years old at conception. As a preliminary step, we also check in this case whether there is any discontinuity in the distribution of age around

the age-35 threshold. We perform a number of statistical manipulation tests (Cattaneo et al. 2018), and find no statistical evidence of systematic discontinuity in the density of mothers' age at conception.[23] At the bottom of Table 5, we present estimation results for equation (1) where the treatment assignment dummy variable A_i is entered both linearly and interacted with the binary indicator *Age Group 35+* (see Lalive 2008 for a similar exercise). The main finding is that prenatal test take-up increases after the policy; this is mostly explained by a higher probability of undergoing a screening test (columns 1–4 of Table 5, coefficients for the *Treatment Assignment* A_i variable). We do not observe any statistically significant difference in the policy effects for women older than 35, as the interaction term is never statistically significant in all columns of Table 5. Finally, we observe that when a woman turns 35, the probability of a screening test decreases, while the probability of a diagnostic test increases (by 12–13.5 percentage points, columns 5–6 of Table 5, coefficient for the binary variable *Age Group 35+*). We do not find evidence of substitution effects as the elimination of the co-payment did not affect the utilization for the riskier diagnostic tests.

In Table 6 we further investigate whether mothers with different educational backgrounds respond differently to the elimination of the co-payment. We find that the probability of women with a low level of education undergoing prenatal tests increases by 7 to 10 percentage points (columns 1–2 of Table 6). For women with a medium level of education, we find a significant increase of 8 to 12 percentage points. For women with high levels of education, we find an increase by only 5 to 6 percentage points (though only the linear specification is statistically significant). We find a statistically significant increase in the take-up rate of noninvasive screening tests only for women with a medium level of education (7 to 10 percentage points). The increase in the probability of undergoing a screening test is lower (from 5 to 6 percentage points) but imprecisely estimated for both low- and high-education groups, possibly because of the smaller samples. Overall, most of the impact of the elimination of the co-payment is among women with a medium level of education.

Table 6 also considers heterogeneity in women's responses according to nationality and area of residence. We find that non-natives have a significantly higher response rate for noninvasive screening tests (9 to 15 percentage points) relative to Italians (6 to 7 percentage points). Similarly, women in metropolitan areas have a significantly higher response rate for screening tests, by 12 to 16 percentage points, while there is no significant response by women in non-metropolitan areas. These results are important, as the elimination of the co-payment was justified on *equity* grounds in addition

[23] Figure A1 – Panel B in the Appendix also shows the density distribution (histogram) of women's ages around the cutoff. This is not surprising as the presumed conception date is only partially based on a woman's report. The definitive estimated conception date is defined by the physician on the basis of a number of fetus measures at the first ultrasound exam. This estimation is accurate as it allows for the calculation of the "gestational age," which is a fundamental parameter for a safe and healthy pregnancy (Verloove et al. 1986). Its manipulation would be considered serious medical negligence.

to *efficiency* grounds. The larger response for foreign women and those living in a metropolitan area (characterized by lower utilization rates pre-policy, see Table 3) is fully consistent with the intentions of the policy.

We finally examine whether the policy had a persistent impact given that the co-payment elimination remained in effect in the following years. We re-estimate equation (1) by using two subsequent years as treatment assignment periods (Ito 2015). To estimate the policy effect in one year from its implementation, we substitute the original post-policy data with the sample of women becoming pregnant between October 2010 and October 2011 and we define A_i, the corresponding treatment assignment variable, to be equal to one in this specific post-policy period, and zero otherwise. Similarly, for the effect two years from the implementation, we include data from October 2011 to October 2012, and we redefine the treatment assignment variable A_i. Table 7 shows the results. When we consider the period of October 2010–October 2011 as the post-policy period, we find that the results are similar to those in Table 4 and Table A4: after the elimination of the co-payment, there is a reduction in the rate of those who undergo no tests (6.5 to 7.5 percentage points, columns 1–2 in Table 7) and an increase in the rate of screening tests (8 to 10 percentage points, columns 3–4 in Table 7).

Results for the post-policy period of October 2011–October 2012 are larger and highly significant: the take-up rate of any prenatal test increases by 11–13 percentage points, and correspondingly, the probability of undergoing a screening test increases by 12–14 percentage points. We do not find any effects on the probability of undergoing a diagnostic test (columns 5–6 of Table 7). This evidence indicates that the co-payment elimination had persistent effects, inducing more women to take prenatal screening tests over time. There is no evidence of statistically significant substitution effects with diagnostic tests even in the long run.

6.4 Effects on women's and newborn health

As a final step, we turn to the analysis of the effects of prenatal tests on women's and newborn health outcomes. Given that the co-payment reform had an effect especially on the take-up of screening tests, we focus our analysis on the impact of this test on health outcomes. Given the endogenous nature of the utilization decision, we adopt a 2SLS approach, where the "Screening Test" dummy is instrumented by the policy threshold considering a fuzzy RDD specification, as described in Section 5.2. Results are reported in Tables 8 and 9, for women's and newborns' health outcomes respectively.

OLS estimates for the coefficient of the "Screening Test" dummy are shown at the bottom of Tables 8 and 9. When we do not account for self-selection into prenatal care, we find significant positive effects on folic acid supplements and significant negative effects on alcohol consumption during

pregnancy as far as women's behavior is concerned, and significant negative effects of the screening test on almost all newborn health outcomes (except for newborn length and stillbirth).

The policy threshold is a strong predictor of the take-up decision (as we already know from the estimation results of equation (1) in previous sections), and it is clearly exogenous with respect to health outcomes. The *F*-test on the excluded instrument in the first stage (reported at the bottom of Tables 8 and 9) is always larger than 30.

When controlling for endogeneity, we find an upward bias in the OLS coefficients in the mother equations, and a downward bias in the OLS coefficients in the newborn equations. Considering Table 8 first, we find that screening tests significantly reduce mothers' weight gain during pregnancy, their alcohol consumption, and the probability of hospitalization during pregnancy (columns 1, 4, and 6 of Table 8). In particular, undergoing a screening test reduces weight gain during pregnancy by 4 kg (roughly 8.8 pounds), about one standard deviation. Screening tests also reduce the probability of drinking during pregnancy by 2.5 percentage points when evaluated at the sample mean of the pre-policy period (Table 2 – Panel A). Screening tests also reduce the probability of any hospital admissions during the pregnancy by 1.7 percentage points when evaluated at the sample mean value of the pre-policy period (Table 2 – Panel A). [24] We do not find any significant effect of prenatal tests on the probability of weight gain larger than 15 kg, on smoking, and on folic acid supplements during pregnancy (columns 2, 3, and 5 of Table 8). Once we account for the self-selection of women, our results show that the take-up of a screening test triggers higher attention to nutrition by prospective mothers. At least for the impact on weight gain and hospitalizations, our results appear to be consistent with previous evidence on the benefits of prenatal care provided in the United States within the Medicaid system (Conway and Kutinova 2006).

Table 9 presents the 2SLS results for equation (2) on newborn health, measured by low birth weight, newborn length, and head circumference (columns 1, 2, and 3, respectively); low Apgar Scores (columns 4 and 5); resuscitation (column 6); preterm (column 7), and stillbirth (column 8). We do not find any causal effect of the prenatal screening test take-up decision on newborn health, confirming the results from previous studies carried out for other healthcare systems with different institutional settings (e.g. Shurtz et al. 2016, for Israel; Plachinski 2017 and Sandner et al. 2018, for the United States).

[24] For the the 'Screening test' coefficients, Tables 8 and 9 report the original p-values and the "adjusted" p-values (the sharpened False Discovery Rate, FDR, q-values by Anderson, 2008), to account for multiple testing hypothesis. Given the large number of measured outcomes (14), there are some concerns about multiple inference, as significant coefficients may emerge simply by chance, even if there are no treatment effects. The only outcome that is significant, even after accounting for the conservative multiple hypothesis testing adjustment, is the hospitalization during pregnancy.

We assess the robustness of the results on the subsamples of women based on education, nationality, and residence (figures A4 and A5 in the Appendix). We find large and significant effects of screening tests among women with a medium level of education, who are native born, and who live in metropolitan areas. The largest weight gain reduction is found in the subsample of women with a medium level of education. The screening test reduces the probability of consuming alcohol during pregnancy, especially among Italian-born women. The reduction in the probability of hospitalization during pregnancy is particularly large again among Italian-born women, mothers with a medium level of education, and women living in metropolitan areas. We do not find any effect of the prenatal screening tests on newborns' health in any subsample (Figure A5 in the Appendix).

Finally, one important outcome for women's and newborn health is whether the increase in the take-up of screening tests did affect also the number of voluntary terminations of pregnancy (VTP). Screening tests – whether positive – require to take diagnostic tests to diagnose the presence of chromosomal anomalies; however, we do not find evidence of an impact of the co-payment reform on this second type of tests. In addition, the literature finds evidence of a decline in the likelihood to terminate pregnancy after a diagnosis of genetic disorder over the years. Natoli et al. (2012) notice that termination rates following a prenatal diagnosis of Down syndrome have been decreasing in recent years in the US. Looking at Scottish data, Jacobs et al. (2016) observe that advances in prenatal screening have improved detection rates for aneuploidy (including trisomy), and this has been accompanied by a reduction in termination rates. Unfortunately, our data do not allow a formal econometric testing of this issue. However, we collected monthly data on the number of all VTP, that occurred during the second trimester of pregnancy, within the administrative borders of Piedmont, between July 2008 and October 2010. The source of data is the Italian Statistical Office (Istat). We also compute the abortion ratio equal to the number of VTP to total number of births. Figure 2 shows the fit of a nonparametric polynomial regression model, separately estimated on both sides of the cutoff point (the policy change, in October 2009), and we are unable to find an impact of the reform.

7 Conclusions

Using an RDD framework, this study evaluates the effects of a policy that in 2009 eliminated the co-payment for noninvasive prenatal screening tests in a large Italian region. We make four key findings. First, eliminating the co-payment triggered an economically and statistically significant increase in the take-up rates of prenatal tests by 7-10 percentage points, and such increase is persistent over time.

Second, we do not find any substitution effect with the more expensive and invasive diagnostic tests. The absence of substitution effects is mostly explained by the group of women older than 35 years at conception, who are at higher risk of congenital disorders, and who are thus more likely to undergo an invasive diagnostic test. At the time of the new policy, this group was already exempted from the co-payment for the invasive test and is therefore likely to have been less sensitive to price changes.

Third, we find that the effect of the policy is larger for younger women, foreigners, those residing in metropolitan areas, and those with low and medium levels of education. Hence, the policy change produced the expected effects of increasing the take-up for more disadvantaged groups that had lower screening test rates before the co-payment was eliminated. Finally, we do not find any effect on newborn health outcomes, but we find that screening tests positively affect mothers' health behaviors as measured by less weight gain, less alcohol consumption, and fewer hospital admissions during pregnancy.

In terms of policy implications, our analysis suggests that eliminating co-payments can be an effective policy lever to encourage the take-up of screening tests. This helps to address both *efficiency* concerns, due to relevant market failures, such as the underestimation of benefits associated with the use of prenatal care, and *equity* concerns, as the increase in prenatal care access is higher among more-disadvantaged groups. It is also worth highlighting that the cost of eliminating the co-payment was relatively inexpensive.[25]

In terms of allocative efficiency, eliminating the co-payment for noninvasive screening tests does not lead to any substitution effect between noninvasive and invasive tests. If the rate of invasive tests is deemed to be too high by policymakers due to excessive costs and the high number of miscarriages, other policy instruments will have to be considered. For instance, a woman's access to an invasive diagnostic test might become conditional on the presence of a high risk of genetic disorders, assessed on the basis of positive results of a previous screening test rather than on the mother's age alone.

Finally, our analysis suggests that increasing the take-up of noninvasive screening tests has a positive impact on maternal health. Improved nutrition and better access to *other* prenatal care services are possible channels for explaining these health benefits. Therefore, screening tests appear to generate positive spillover toward other policy efforts aimed at improving maternal health. We lack additional data (e.g. on nutrition or other prenatal care services utilization) to rule out (or validate) these explanations. Future research could explore those efforts.

[25] The resolution of the Regional Government (Delibere della Giunta Regionale) n. 38-11960, August 4, 2009 established that the annual additional costs for the public administration following the co-payment elimination was €500,000. The annual health expenditure of the Regional Government is around €8 billion.

8 Acknowledgements

The authors acknowledge Sonia Bhalotra, Janet Currie, Osea Giuntella, Jo Thori Lind, Elena Lucchese, Gianmaria Martini, Catia Nicodemo, Anne-Fleur Roos, and Francesca Zantomio for their insightful comments on previous drafts. We are grateful to the participants at the 2019 Healthcare Management & Economics Workshop (HEMAW), the 6th Health Econometrics Workshop (HEW), the 2018 Annual Meeting of the European Public Choice Society (EPCS), the XXX Annual Conference of the Italian Society of Public Economics (SIEP). We also thank Maria Maspoli and Gloria Prina at the Assessorato Regionale alla Sanità, Regione Piemonte for help with the data. All remaining errors are our own.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations of interest: none.

9 References

Almond, D., K.Y. Chay, and D.S. Lee. 2005. The costs of low birth weight. *Quarterly Journal of Economics*, Vol. 120(3), p. 1031–1083.

Almond, D., and J. Currie. 2011. Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*, Vol. 25(3), p. 153-72.

Anderson, M.L. 2008. Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, Vol. 103(484), p. 1481-1495.

Angrist, J., and J-S Pishke. 2009. Mostly Harmless Econometrics. Princeton University Press.

AReSS (Agenzia Regionale per i Servizi Sanitari). 2008. Tavolo A.Re.S.S. sullo Screening delle Anomalie Fetali Proposta di riorganizzazione del percorso per lo screening delle anomalie fetali. Technical Report.

Barreca, A., M. Guldi, J. Lindo, and G. Waddell. 2011. Saving Babies? Revisiting the effect of very low birth weight classification. *Quarterly Journal of Economics*, Vol. 126(4), p. 2117-2123.

Bitler, M. P., and C. S. Carpenter. 2016. Health Insurance Mandates, Mammography, and Breast Cancer Diagnoses. *American Economic Journal: Economic Policy,* Vol. 8(3), p. 39-68.

Bitler, M. P., and C. S. Carpenter. 2017. Effects of State Cervical Cancer Insurance Mandates on Pap Test Rates. *Health Services Research*, Vol. 52(1), p. 156-175.

Borra, C., L. González, and A. Sevilla. 2016. Birth Timing and Neonatal Health. *American Economic Review*, Vol. 106 (5), p. 329-32.

Boyd, P., C. DeVigan, B. Khoshnood, M. Loane, E. Garne, and H. Dolk. 2008. Survey of prenatal screening policies in Europe for structural malformations and chromosome anomalies, and their impact on detection and termination rates for neural tube defects and downs syndrome. *BJOG: An International Journal of Obstetrics & Gynaecology*, Vol. 115(6), p. 689-696.

Calonico, S., M.D. Cattaneo, and R. Titiunik. 2014. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs, *Econometrica*, Vol. 82(6), p. 2295-2326.

Calonico, S., M.D. Cattaneo, and R. Titiunik. 2015. Optimal Data-Driven Regression Discontinuity Plots, *Journal of the American Statistical Association*, Vol. 110(512), p. 1753-1769.

Calonico, S., M.D. Cattaneo, and M.H. Farrell. 2018. On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference. *Journal of the American Statistical Association*, Vol. 113(522), p. 767-779.

Cattaneo, M.D., M. Jansson, and X. Ma. 2018. Manipulation Testing based on Density Discontinuity. *Stata Journal*, Vol. 18(1), p. 234-261.

Cohen, J., P. Dupas, and S. Schaner. 2015. Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial. *American Economic Review*, Vol. 105(2), p. 609-645.

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

Corman, H., D.M. Dave, and N. Reichman. 2018. Effects of Prenatal Care on Birth Outcomes: Reconciling a Messy Literature. NBER Working Paper No. 24885.

Crombag N., Y. Vellinga, S. Kluijfhout, L. Bryant, P. Ward, R. Iedema-Kuiper, P. Schielen, J. Bensing, G. Visser, A. Tabor, and J. Hirst. 2014. Explaining variation in Down's syndrome screening uptake: comparing the Netherlands with England and Denmark using documentary analysis and expert stakeholder interviews, *BMC Health Services Research*, Vol. 14, p. 437.

Crombag, N. 2016. Explaining low uptake for Down syndrome screening in the Netherlands (and predicting utilisation of other programmes). Mimeo, Utrecht University, the Netherlands.

Currie, J. 2006. The Take-up of Social Benefits. In A. Auerbach, D.Card, and J. Quigley (eds), "*Poverty, the Distribution of Income, and Public Policy*," New York: Russell Sage, p. 80-148.

Currie, J., S. DellaVigna, E. Moretti, and V. Pathania. 2010. The Effect of Fast Food Restaurants on Obesity and Weight Gain. *American Economic Journal: Economic Policy*, Vol. 2(3), p. 32-63.

Currie, J., and M. Rossin-Slater. 2015. Early-Life Origins of Life-Cycle Well-Being: Research and Policy Implications. *Journal of Policy Analysis and Management*, Vol. 36(4), p. 974-974.

Dormandy, E., S. Michie, R. Hooper, and T. Marteau. 2005. Low uptake of prenatal screening for Down syndrome in minority ethnic groups and socially deprived groups: a reflection of women's attitudes or a failure to facilitate informed choices? *International Journal of Epidemiology*, Vol. 34, p. 346–352.

EUROCAT. 2010. Special Report: Prenatal Screening Policies in Europe 2010. EUROCAT (European Surveillance of Congenital Anomalies) Central Registry, University of Ulster.

Fajnzylber, E., V.J. Hotz, and S.G. Sanders. 2010. An economic model of amniocentesis choice, NBER Working Paper 16306.

Figlio, D., J. Guryan, K. Karbownik, and J. Roth. 2014. The Effects of Poor Neonatal Health on Children's Cognitive Development. *American Economic Review*, Vol. 104(12), p. 3921–3955.

Fransen, M.P., M.H. Schoonen, J.P Mackenbach, E.A. Steegers, H.J. de Koning, J.A. Laudy, R.J. Galjaard, C.W. Looman, M.L. Essink-Bot, and H.I. Wildschut. 2010. Ethnic differences in participation in prenatal screening for Down syndrome: A register-based study. *Prenatal Diagnosis*, Vol. 30, p. 988–994.

Gajdos, T., C. Garrouste, and P. Geoffard. 2016. The subjective value of a life with Down syndrome: Evidence from amniocentesis decision. *Journal of Economic Behavior & Organization*, Vol. 127, p. 59-69.

Garrouste, C., J. Le, and E. Maurin. 2011. The choice of detecting Down syndrome: Does money matter? *Health Economics*, Vol. 20, p. 1073–1089.

Grimes, D., and K. Schulz. 2002. Uses and abuses of screening tests. *Lancet*, Vol. 359(9309), p. 881 – 884.

Halla M., G.J. Pruckner, and T. Schober. 2016. Cost savings of developmental screenings: Evidence from a nationwide program. *Journal of Health Economics*, Vol. 49, 120-135.

Hausman, C. and D. Rapson. 2018. Regression Discontinuity in Time: Considerations for Empirical Applications. *Annual Review of Resource Economics*, Vol. 10, p. 533-552.

Imbens, G., and T. Lemieux. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, Vol. 142, p. 615–635.

Institute of Medicine. 2009. Weight Gain During Pregnancy: Re-examining the Guidelines. Report Brief.

Ito, K. 2015. Asymmetric Incentives in Subsidies: Evidence from a Large-Scale Electricity Rebate Program. *American Economic Journal: Economic Policy*, Vol. 7(3), p. 209-237.

Jacobs M., S.A. Cooper, R. McGowan, S.M. Nelson, and J.P. Pell JP. 2016. Pregnancy Outcome following Prenatal Diagnosis of Chromosomal Anomaly: A Record Linkage Study of 26,261 Pregnancies. *PLoS One*, Vol. 11(12), e0166909.

Javaher, P., E. Nyoungui, H. Kääriäinen, U. Kristoffersson, I. Nippert, J. Sequeiros, and J. Schmidtke. 2010. Genetic Screening in Europe. *Public Health Genomics*, Vol. 13, p. 524–537.

Kontopantelis, E., T. Doran, D.A. Springate, I. Buchan, and D. Reeves. 2015. Regression based quasi-experimental approach when randomisation is not an option: Interrupted time series analysis. *British Medical Journal*, Vol. 350, h.2750.

Lalive R. 2008. How do extended benefits affect unemployment duration? A regression discontinuity approach. *Journal of Econometrics*, Vol. 142, p. 785–806.

Lee D., and D. Card. 2008. Regression discontinuity inference with specification error. *Journal of Econometrics*, Vol. 142, p. 655-674.

Lee D., and T. Lemieux. 2010. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2), p. 281-355.

Loane M., J. Morris, M. Addor, L. Arriola, J. Budd, B. Doray, E. Garne, M. Gatt, M. Haeusler, and B. Khoshnood. 2013. Twenty-year trends in the prevalence of Down syndrome and other trisomies in Europe: Impact of maternal age and prenatal screening. *European Journal of Human Genetics*, 21(1), p. 27-33.

Malone, F.D., J. Canick, R.H. Ball, D.A. Nyberg, C.H. Comstock, R. Bukowski, R.L. Berkowitz, S.J. Gross, L. Dugoff, S.D. Craigo, I.E. Timor-Tritsch, and S.R. Carr. 2005. First-Trimester or Second-Trimester Screening, or Both, for Down's Syndrome. *New England Journal of Medicine*, Vol. 353(19), p. 2001-2011.

McCrary, J. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, Vol. 142(2), p. 698-714.

Metcalfe, A., L.M. Lix, J. Johnson, F. Bernier, G. Currie, A.W. Lyon, and S.C. Tough. 2013. Assessing the Impact of the SOGC Recommendations to Increase Access to prenatal Screening on Overall Use of Health Resources in Pregnancy. *Journal of Obstetrics and Gynaecology Canada*, Vol. 35 (5), p. 444-453.

Natoli, J. L., D.L. Ackerman, S. Mcdermott, S., and J.G. Edwards. 2012. Prenatal diagnosis of Down syndrome: A systematic review of termination rates (1995-2011). *Prenatal Diagnosis*, Vol. 32(2), p. 142–153.

Palomaki, G.E., K. Steinort, G.J. Knight, and J.E. Haddow. 2006. Comparing three screening strategies for combining first- and second-trimester Down syndrome markers. *Obstetrics & Gynecology*, Vol. 107, p. 367-75.

Palomaki, G.E., G.J. Knight, E.R. Ashwood, R.G. Best, J.E. Haddow. 2013. Screening for down syndrome in the United States: Results of surveys in 2011 and 2012. *Archives of Pathology & Laboratory Medicine.* Vol. 137, p. 921–926.

Plachinski, L. 2017. The Effect of Access to Prenatal Genetic Testing on Test Utilization and Birth Outcomes: Evidence from Down syndrome. Wellesley College, Honors Thesis Collection n. 475.

Sandner, M., T. Cornelissen, T. Jungmann, and P. Herrmann. 2018. Evaluating the effects of a targeted home visiting program on maternal and child health outcomes. *Journal of Health Economics*, Vol. 58, p. 269-283.

Santalahti, P., E. Hemminki, A.-M. Latikka, and M. Ryynänen. 1998. Women's decision-making in prenatal screening. *Social Science and Medicine*, Vol. 46(8), p. 1067-1076.

Seror, V. 2008. Fitting Observed and Theoretical Choices--Women's Choices about Prenatal Diagnosis of Down Syndrome. *Health Economics*, Vol. 17(5), p. 557-77.

Shurtz, I., A. Brzezinski, and A. Frumkinc. 2016. The impact of financing of screening tests on utilization and outcomes: The case of amniocentesis. *Journal of Health Economics*, Vol. 48, p. 61-73.

Turati, G. 2013. The Italian Servizio Sanitario Nazionale: a renewing tale of lost promises. In J. Costa-Font, S.C. Greer (Eds.), *Federalism and Decentralization in European Health and Social Care: Competition, Innovation, and Cohesion*, Palgrave MacMillan, London, p. 47-66.

Vassy, C., S. Rosman, and B. Rousseau. 2014. From policy making to service use. Down's syndrome antenatal screening in England, France and the Netherlands. *Social Science & Medicine*, Vol. 106, p. 67-74.

Verloove, P., R.A. Verwey, R. Brand, and M.J. Keirse. 1986. Importance of gestational age. *Lancet,* Vol. 1(8496), p. 1494.

Wald, N. J., C. Rodeck, A. K. Hackshaw, J. Walters, L. Chitty, and A.M. Mackinson - SURUSS Research Group. 2003. First and second trimester antenatal screening for Down's syndrome: the results of the serum, urine and ultrasound screening study (SURUSS). *Health Technology Assessment*, Vol. 7(11), p. 1–77.

Woodhouse, C., J. Lopez Camelo, and G.L. Wehby. 2014. A comparative analysis of prenatal care and fetal growth in eight South American countries. *PLoS One*, Vol. 9(3), p. e91292.

World Health Organization. 2016. *WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience*. World Health Organization Report.

Yan, J. 2017. The Effects of Prenatal Care Utilization on Maternal Health and Health Behaviors. *Health Economics*, Vol. 26(8), p. 1001-1018.

Table 1. Number and Proportion of Women Undergoing Different Prenatal Tests

	Whole Sample	Pre-Policy	Post-Policy
No Test	7,664	4,791	2,873
	23.35	28.09	18.21
Screening Test Only	21,848	10,609	11,239
	66.55	62.21	71.24
Diagnostic Test Only	2,297	1,139	1,158
	7.00	6.68	7.34
Both Tests	1,020	514	506
	3.10	3.02	3.21
Total	32,829	17,053	15,776
	100	100	100

Notes: The first row shows *frequencies* and the second row shows *column percentages*.

Table 2. Summary Statistics of Health Outcomes.

Panel A. Mothers' health outcomes.

	Whole Sample			Pre-Policy		Post-Policy	
	Mean	***SD***	***N***	***Mean***	***SD***	***Mean***	***SD***
Weight Gain in Pregnancy	13.125	4.673	27,874	13.218	4.870	13.033	4.467
Weight Gain in Preg. >15 kg	0.240	0.427	27,874	0.245	0.430	0.235	0.424
Smoke in Pregnancy	0.073	0.260	32,829	0.071	0.257	0.075	0.264
Alcohol in Pregnancy	0.041	0.199	32,829	0.054	0.225	0.028	0.164
Folic Acid in Pregnancy	0.826	0.379	32,829	0.794	0.404	0.860	0.347
Hospital Admissions in Preg.	0.037	0.189	22,935	0.035	0.185	0.039	0.193

Panel B. Newborn health outcomes.

	Whole Sample			Pre-Policy		Post-Policy	
	Mean	***SD***	***N***	***Mean***	***SD***	***Mean***	***SD***
Low Weight	0.055	0.229	33,048	0.054	0.225	0.057	0.232
Newborn Length	49.51	2.010	32,333	49.55	1.996	49.47	2.024
Head Circumference	34.00	1.31	32,427	34.01	1.299	33.98	1.321
Low Apgar-1 min	0.202	0.401	32,118	0.204	0.403	0.200	0.400
Low Apgar-5 min	0.055	0.227	32,135	0.056	0.230	0.053	0.224
Resuscitation	0.034	0.182	33,049	0.034	0.181	0.034	0.182
Preterm	0.054	0.227	33,049	0.054	0.227	0.054	0.226
Stillbirth	0.003	0.051	33,049	0.003	0.051	0.003	0.050

Notes: All variables are binary variables, except for "Newborn Length" and "Head Circumference" both expressed in centimeters (cm), and "Weight Gain in Pregnancy" expressed in kilograms (kg).

Table 3. Percentage of Women Who *Do Not Undergo Any* Prenatal Tests, Within Different Subsamples

	Whole Sample	Pre-Policy	Post-Policy
Age groups:			
Age group 18–24	38.5	43.4	33.2
Age group 25–29	24.0	28.6	19.2
Age group 30–34	19.0	23.6	13.8
Age group 35+	17.5	22.8	12.1
Education levels:			
Low Education	31.7	36.8	25.9
Medium Education	21.0	25.4	16.1
High education	17.9	22.2	13.9
Nationality:			
Foreign Born	38.7	43.5	33.7
Native	18.7	23.5	13.3
Area of residence:			
Nonmetropolitan area	21.7	25.6	17.2
Metropolitan area	25.9	32.2	19.7

Table 4. Effect of Policy Change on the Prenatal Tests' Take-Up Rates

Dep. Var.:	*No Test*	*Screening Test*	*Diagnostic Test*	*Both Tests*
	(1)	(2)	(3)	(4)
Assignment A_i	-0.073	0.065	0.008	-0.0002
	(0.03)	(0.03)	(0.01)	(0.005)
Linear Trend, (C_i-C*)	-0.001	0.001	0.0001	0.00007
	(0.00)	(0.00)	(0.0002)	(0.0001)
Assignment A_i x Linear Trend, (C_i-C*)	0.001	-0.001	-0.0004	-0.0001
	(0.00)	(0.00)	(0.0003)	(0.0002)
Age group 25–29	-0.067	0.073	-0.005	-0.001
	(0.01)	(0.01)	(0.002)	(0.002)
Age group 30–34	-0.089	0.079	0.003	0.007
	(0.01)	(0.01)	(0.003)	(0.002)
Age group 35+	-0.093	-0.239	0.269	0.064
	(0.02)	(0.02)	(0.01)	(0.004)
Medium Education	-0.026	0.018	0.003	0.005
	(0.01)	(0.01)	(0.003)	(0.002)
High education	-0.027	0.018	0.005	0.004
	(0.01)	(0.01)	(0.005)	(0.002)
Employed	-0.080	0.069	0.007	0.004
	(0.01)	(0.01)	(0.003)	(0.003)
Married	0.036	-0.014	-0.019	-0.003
	(0.01)	(0.01)	(0.003)	(0.00)
Native	-0.111	0.091	0.014	0.006
	(0.01)	(0.01)	(0.003)	(0.002)
Father Employed	-0.176	0.147	0.022	0.008
	(0.04)	(0.03)	(0.01)	(0.004)
Twin	0.017	-0.004	-0.001	-0.012
	(0.02)	(0.02)	(0.02)	(0.01)
Past Miscarriage	-0.017	0.014	-0.001	0.005
	(0.01)	(0.01)	(0.01)	(0.003)
Past Abortion	-0.036	0.027	0.008	0.001
	(0.01)	(0.01)	(0.01)	(0.004)
Metropolitan Area	0.030	-0.026	-0.005	0.0002
	(0.01)	(0.01)	(0.003)	(0.002)
Constant	0.614	0.396	-0.010	0.001
	(0.04)	(0.04)	(0.01)	(0.005)
Adj. R^2	0.09	0.10	0.20	0.02
Mean Dep. Var.	0.233	0.666	0.070	0.031
N. Obs.	32,829	32,829	32,829	32,829

Notes: Each column presents the estimation of equation (1) by OLS, on the sample of women becoming pregnant during the 52 weeks before and the 52 weeks after the policy change. The dependent variables are: No Test (equal to one if the woman had no prenatal tests, and zero otherwise) in column 1; Screening Test (equal to one if a screening test is undertaken, and zero otherwise) in column 2; Diagnostic Test (equal to one if an invasive diagnostic test is undertaken, and zero otherwise) in column 3; and Both Tests (equal to one if the woman had a screening test followed by an invasive diagnostic test, and zero otherwise) in column 4. Assignment A_i, is equal to one in the post- policy period. All equations include the term $(C_i - C^*)$, that is, the number of weeks between the conception date C_i, and the cutoff date C^*, and its interaction with the treatment assignment variable. Standard errors (in parentheses) are clustered at the level of district of residence–quarter of conception (138 clusters).

Table 5. Effect of Policy Change on the Prenatal Tests' Take-Up Rates: Age Groups

Dep. Var.:	***No Test***		***Screening Test***		***Diagnostic Test***		***Both Tests***	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Specification:	*Linear*	*Quadratic*	*Linear*	*Quadratic*	*Linear*	*Quadratic*	*Linear*	*Quadratic*
1. Subsample: Age 18–24 *(N. Obs. 5,424)*								
Assignment A_i	-0.074	-0.108	0.072	0.113	0.004	0.003	-0.002	-0.007
	(0.05)	(0.08)	(0.05)	(0.08)	(0.004)	(0.01)	(0.01)	(0.01)
Adj. R^2	0.08	0.08	0.07	0.07	0.01	0.01	0.003	0.003
Mean Dep. Var.	0.385		0.599		0.005		0.011	
2. Subsample: Age 25–29 *(N. Obs. 9,280)*								
Assignment A_i	-0.075	-0.100	0.077	0.094	0.001	0.007	-0.002	-0.002
	(0.04)	(0.06)	(0.04)	(0.06)	(0.003)	(0.005)	(0.005)	(0.01)
Adj. R^2	0.06	0.06	0.05	0.05	0.002	0.003	0.002	0.002
Mean Dep. Var.	0.240		0.738		0.008		0.014	
3. Subsample: Age 30–34 *(N. Obs. 11,196)*								
Assignment A_i	-0.072	-0.101	0.072	0.105	-0.002	-0.006	0.002	0.001
	(0.03)	(0.04)	(0.03)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)
Adj. R^2	0.07	0.07	0.05	0.05	0.01	0.01	0.003	0.003
Mean Dep. Var.	0.190		0.767		0.019		0.024	
4. Subsample: Age 35+ *(N. Obs. 6,929)*								
Assignment A_i	-0.079	-0.086	0.055	0.041	0.023	0.039	-0.0002	0.006
	(0.04)	(0.05)	(0.04)	(0.05)	(0.03)	(0.05)	(0.02)	(0.02)
Adj. R^2	0.07	0.07	0.07	0.07	0.06	0.06	0.01	0.01
Mean Dep. Var.	0.175		0.457		0.286		0.082	
5. Subsample: Age 32–38 *(N. Obs. 10,102)*								
Assignment A_i	-0.078	-0.094	0.077	0.090	-0.001	-0.003	0.001	0.007
	(0.04)	(0.05)	(0.03)	(0.05)	(0.01)	(0.02)	(0.01)	(0.01)
Age Group 35+	-0.004	0.018	-0.156	-0.158	0.135	0.120	0.024	0.020
	(0.03)	(0.03)	(0.04)	(0.06)	(0.03)	(0.05)	(0.02)	(0.02)
Assignment A_i x Age Group 35+	0.007	0.001	-0.029	-0.019	0.018	0.021	0.005	-0.002
	(0.03)	(0.04)	(0.05)	(0.07)	(0.04)	(0.06)	(0.02)	(0.03)
Adj. R^2	0.08	0.08	0.11	0.11	0.14	0.14	0.01	0.01
Mean Dep. Var.	0.181		0.642		0.128		0.049	

Notes: Each row presents the estimation of equation (1) by OLS, for different subsamples of women. The table reports only the post-policy dummy variable Assignment A_i. All equations include the term $(C_i - C^*)$, its square (only in even columns 2, 4, 6, and 8), and the corresponding interactions with the treatment assignment variable. All equations also include the full set of observable characteristics for women. Standard errors (in parentheses) are clustered at the level of district of residence–quarter of conception.

Table 6. Effect of Policy Change on the Prenatal Tests' Take-Up Rates: Education Level, Nationality, and Residence

Dep. Var.:	***No Test***		***Screening Test***		***Diagnostic Test***		***Both Tests***	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Specification:	*Linear*	*Quadratic*	*Linear*	*Quadratic*	*Linear*	*Quadratic*	*Linear*	*Quadratic*
1. Subsample: Low Education *(N. Obs. 9,277)*								
Assignment A_i	-0.072	-0.099	0.054	0.061	0.014	0.028	0.005	0.009
	(0.04)	(0.05)	(0.04)	(0.05)	(0.01)	(0.01)	(0.01)	(0.01)
Adj. R^2	0.12	0.12	0.10	0.10	0.19	0.19	0.02	0.02
Mean Dep. Var.	0.317		0.614		0.047		0.021	
2. Subsample: Medium Education *(N. Obs. 16,467)*								
Assignment A_i	-0.077	-0.116	0.072	0.104	0.003	0.010	0.001	0.002
	(0.04)	(0.05)	(0.03)	(0.05)	(0.01)	(0.02)	(0.01)	(0.01)
Adj. R^2	0.07	0.07	0.11	0.11	0.20	0.20	0.02	0.02
Mean Dep. Var.	0.210		0.684		0.073		0.033	
3. Subsample: High Education *(N. Obs. 7,085)*								
Assignment A_i	-0.064	-0.048	0.062	0.069	0.010	-0.006	-0.008	-0.015
	(0.03)	(0.05)	(0.03)	(0.05)	(0.02)	(0.03)	(0.01)	(0.01)
Adj. R^2	0.04	0.04	0.09	0.09	0.17	0.17	0.02	0.02
Mean Dep. Var.	0.179		0.689		0.093		0.039	
4. Subsample: Native *(N. Obs. 25,141)*								
Assignment A_i	-0.070	-0.087	0.058	0.068	0.010	0.016	0.002	0.003
	(0.03)	(0.05)	(0.03)	(0.04)	(0.01)	(0.02)	(0.01)	(0.01)
Adj. R^2	0.05	0.05	0.11	0.11	0.20	0.20	0.02	0.02
Mean Dep. Var.	0.187		0.692		0.085		0.036	
5. Subsample: Foreign Born *(N. Obs.7,688)*								
Assignment A_i	-0.078	-0.129	0.086	0.145	-0.000	-0.005	-0.007	-0.011
	(0.05)	(0.06)	(0.05)	(0.06)	(0.01)	(0.01)	(0.01)	(0.01)
Adj. R^2	0.07	0.07	0.06	0.06	0.12	0.12	0.02	0.01
Mean Dep. Var.	0.387		0.578		0.020		0.015	
6. Subsample: Metropolitan *(N. Obs. 12,784)*								
Assignment A_i	-0.142	-0.186	0.121	0.161	0.022	0.023	-0.001	0.003
	(0.06)	(0.07)	(0.05)	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)
Adj. R^2	0.08	0.09	0.09	0.09	0.19	0.19	0.02	0.02
Mean Dep. Var.	0.259		0.638		0.071		0.032	
7. Subsample: Nonmetropolitan *(N. Obs. 20,045)*								
Assignment A_i	-0.027	-0.037	0.027	0.034	-0.002	0.004	0.001	-0.002
	(0.03)	(0.04)	(0.03)	(0.04)	(0.01)	(0.02)	(0.01)	(0.01)
Adj. R^2	0.10	0.10	0.11	0.11	0.20	0.20	0.02	0.02
Mean Dep. Var.	0.217		0.683		0.070		0.030	

Notes: Each row presents the estimation of equation (1) by OLS, for different subsamples of women. The table reports only the post-policy dummy variable Assignment A_i. All equations include the trend $(C_i - C^*)$, its squared term (only in even columns 2, 4, 6, and 8), and the corresponding interactions with the treatment assignment variable. All equations also include the full set of observable characteristics for women. Standard errors (in parentheses) are clustered at the level of district of residence–quarter of conception.

Table 7. Persistence in the Effect of Policy Change on the Prenatal Tests' Take-Up Rates

Dep. Var.:	***No Test***		***Screening Test***		***Diagnostic Test***		***Both Tests***	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Specification:	*Linear*	*Quadratic*	*Linear*	*Quadratic*	*Linear*	*Quadratic*	*Linear*	*Quadratic*
Policy Change on Oct. 1, 2010								
(N. Obs. 31,707)								
Assignment A_i	-0.065	-0.075	0.083	0.102	-0.009	-0.014	-0.010	-0.014
	(0.04)	(0.05)	(0.03)	(0.05)	(0.01)	(0.01)	(0.01)	(0.01)
Adj. R^2	0.10	0.10	0.11	0.11	0.18	0.18	0.03	0.03
Mean Dep. Var.	0.222		0.682		0.061		0.034	
Policy Change on Oct. 1, 2011								
(N. Obs. 30,336)								
Assignment A_i	-0.114	-0.133	0.123	0.142	-0.013	-0.016	0.004	0.007
	(0.03)	(0.05)	(0.03)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)
Adj. R^2	0.10	0.10	0.11	0.11	0.19	0.19	0.03	0.03
Mean Dep. Var.	0.225		0.681		0.063		0.030	

Notes: Each row presents the estimation of equation (1) by OLS, for different simulated samples. The table reports only the post-policy dummy variable Assignment A_i. All equations include the term $(C_i - C^*)$, its square (only in even columns 2, 4, 6, and 8), and the corresponding interactions with the treatment assignment variable. All equations also include the full set of observable characteristics for women. When we simulate policy change on October 1, 2010, we substitute the post-policy data with the data for women who became pregnant between October 2010 and October 2011. When we simulate policy change on October 1, 2011, we substitute the post-policy data with the data for women who became pregnant between October 2011 and October 2012. Standard errors (in parentheses) are clustered at the level of district of residence– quarter of conception (144 clusters for the policy change on October 1, 2010, and 134 clusters for a policy change on October 1, 2011).

Table 8. Effect of the Prenatal Screening Tests on Health Outcomes of the Mother: 2SLS Results

Dep. Var.:	*Weight Gain in Pregnancy*	*Weight Gain in Pregnancy >15 kg*	*Smoke in Pregnancy*	*Alcohol in Pregnancy*	*Folic Acid in Pregnancy*	*Hospital Admissions in Pregnancy*
	(1)	(2)	(3)	(4)	(5)	(6)
Screening Test	-4.141	-0.233	-0.060	-0.469	-0.108	-0.495
	(2.04)	(0.15)	(0.11)	(0.21)	(0.42)	(0.17)
[p-value]	[0.04]	[0.12]	[0.58]	[0.03]	[0.79]	[0.004]
{Sharpened FDR q-value}	{0.23}	{0.28}	{0.96}	{0.21}	{0.99}	{0.06}
Linear Trend, (C_i-C*)	0.002	0.0003	0.0003	0.001	0.002	0.002
	(0.003)	(0.0003)	(0.0003)	(0.001)	(0.001)	(0.001)
Assignment A_i x Linear Trend (C_i-C*)	0.001	-0.0004	-0.0001	-0.001	-0.002	-0.002
	(0.01)	(0.0005)	(0.0003)	(0.001)	(0.001)	(0.001)
Age group 25–29	0.302	0.013	-0.012	0.039	0.033	0.030
	(0.21)	(0.01)	(0.01)	(0.02)	(0.03)	(0.01)
Age group 30–34	-0.057	-0.018	-0.024	0.055	0.032	0.036
	(0.26)	(0.02)	(0.01)	(0.02)	(0.04)	(0.02)
Age group 35+	-1.734	-0.126	-0.047	-0.063	-0.010	-0.100
	(0.33)	(0.03)	(0.02)	(0.04)	(0.07)	(0.03)
Medium Education	-0.114	-0.018	-0.029	0.020	0.041	0.007
	(0.08)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
High education	-0.639	-0.073	-0.059	0.017	0.055	0.003
	(0.10)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Employed	0.677	0.030	-0.001	0.038	0.053	0.026
	(0.16)	(0.01)	(0.01)	(0.02)	(0.03)	(0.01)
Married	-0.560	-0.051	-0.054	-0.009	0.011	-0.001
	(0.08)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
Native	0.322	0.004	0.061	0.046	0.110	0.051
	(0.24)	(0.02)	(0.01)	(0.02)	(0.04)	(0.02)
Father Employed	0.708	0.028	0.009	0.079	0.190	0.067
	(0.20)	(0.02)	(0.02)	(0.04)	(0.08)	(0.03)
Twin	2.591	0.252	-0.011	-0.016	0.018	0.062
	(0.31)	(0.03)	(0.01)	(0.02)	(0.02)	(0.02)
Past Miscarriage	0.193	0.017	0.019	0.019	0.074	0.022
	(0.10)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
Past Abortion	0.723	0.066	0.073	0.049	0.048	0.030
	(0.11)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Metropolitan Area	0.039	0.009	0.007	0.044	0.037	0.000
	(0.09)	(0.01)	(0.00)	(0.01)	(0.02)	(0.01)
Constant	15.383	0.431	0.133	0.200	0.553	0.278
	(1.03)	(0.07)	(0.05)	(0.09)	(0.19)	(0.08)
F-Test excluded instrument [*p*-value]	66.83 [0.00]	66.83 [0.00]	46.38 [0.00]	46.38 [0.00]	46.38 [0.00]	47.36 [0.00]
OLS Results:						
Screening Test	0.009	-0.002	0.005	-0.027	0.155	-0.005
	(0.07)	(0.01)	(0.00)	(0.01)	(0.02)	(0.00)
Adj. R^2	0.02	0.02	0.03	0.03	0.11	0.01
N. Obs.	27,874	27,874	32,829	32,829	32,829	22,935

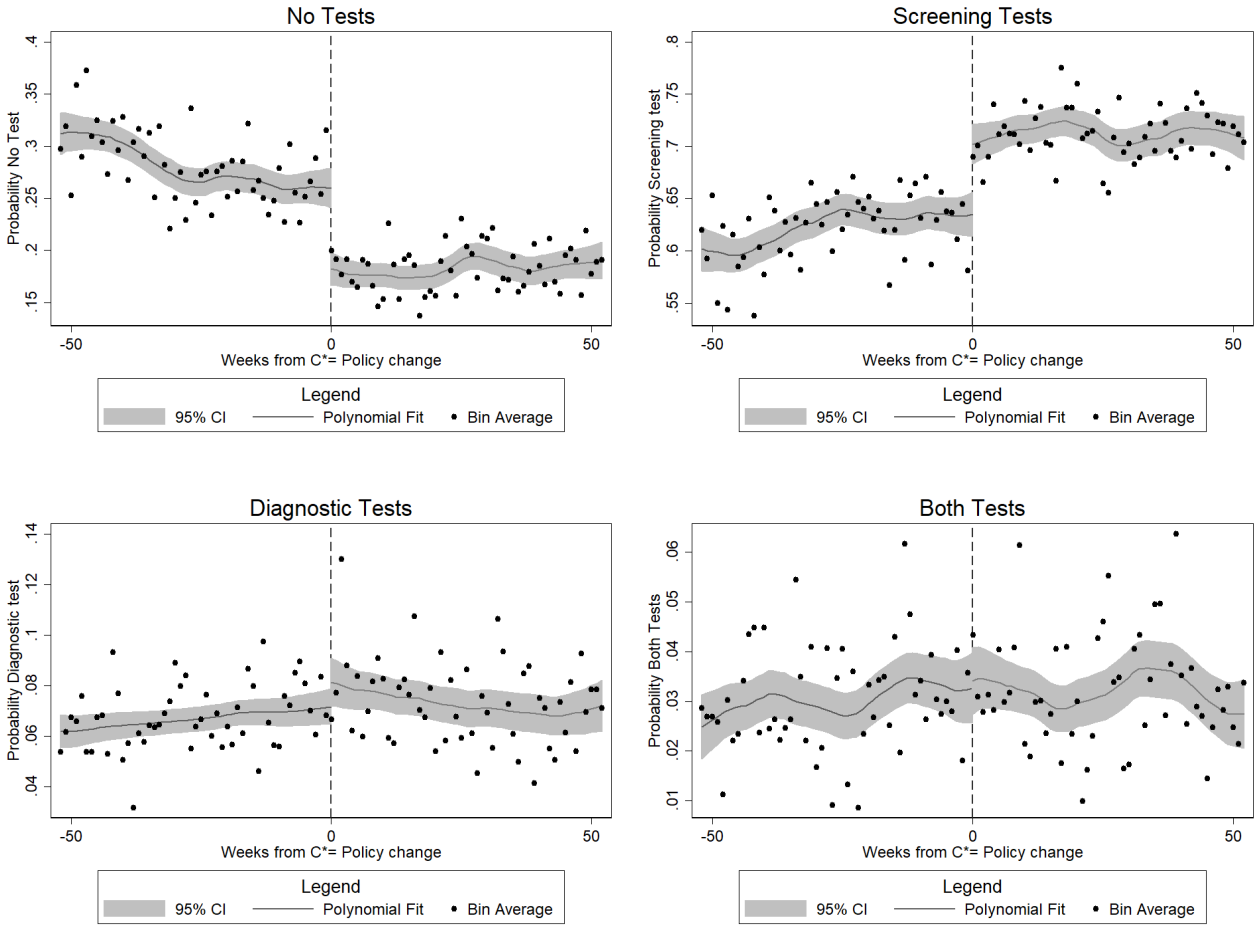
Notes: Each column presents the estimation of equation (2) by 2SLS. The *Screening Test* variable is a binary variable equal to one if a woman had a noninvasive screening test during pregnancy, and zero otherwise. *Assignment A_i*, equal to one in the post-policy period, is the instrumental variable for *Screening Test*. Standard errors (in round parentheses) are clustered at the level of district of residence–quarter of conception. OLS results for equation (2) are reported (Screening Test coefficient and Adj. R^2 only). Below the 'Screening test' coefficient, the table reports the original p-values (square brackets) and the "adjusted" p-values (in braces, the sharpened False Discovery Rate, FDR, q-values by Anderson, 2008), to account for multiple testing hypothesis.

Table 9. Effect of the Prenatal Screening Tests on Health Outcomes of the Newborn: 2SLS Results

Dep. Var.:	*Low Weight*	*Newborn Length*	*Head Circumfer.*	*Low Apgar-1 min*	*Low Apgar-5 min*	*Resuscit.*	*Preterm*	*Stillbirth*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Screening Test	0.073	-3.247	-1.879	0.031	-0.021	-0.040	0.069	-0.008
	(0.09)	(2.00)	(1.16)	(0.35)	(0.11)	(0.09)	(0.09)	(0.02)
[p-value]	[0.39]	[0.10]	[0.10]	[0.93]	[0.85]	[0.67]	[0.43]	[0.63]
{Sharpened FDR q-value}	{0.77}	{0.28}	{0.28}	{0.99}	{0.99}	{0.96}	{0.77}	{0.96}
Linear Trend, (C_i-C*)	0.000	0.002	0.002	-0.000	0.000	0.000	0.000	0.000
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Assignment A_i x Lin. Trend (C_i-C*)	-0.000	0.003	0.002	0.000	-0.000	-0.000	-0.000	-0.000
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age group 25–29	-0.002	0.227	0.169	0.002	0.006	0.007	-0.005	0.002
	(0.01)	(0.17)	(0.10)	(0.03)	(0.01)	(0.01)	(0.01)	(0.00)
Age group 30–34	0.003	0.202	0.173	0.011	0.008	0.010	0.000	0.003
	(0.01)	(0.19)	(0.11)	(0.03)	(0.01)	(0.01)	(0.01)	(0.00)
Age group 35+	0.036	-0.690	-0.312	0.030	0.013	0.007	0.029	0.001
	(0.02)	(0.35)	(0.20)	(0.06)	(0.02)	(0.02)	(0.02)	(0.00)
Medium Education	-0.011	0.173	0.073	-0.025	-0.014	-0.005	-0.010	0.000
	(0.00)	(0.06)	(0.04)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
High education	-0.015	0.271	0.189	-0.026	-0.009	-0.007	-0.013	-0.001
	(0.00)	(0.06)	(0.04)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Employed	-0.012	0.294	0.160	-0.007	-0.000	-0.002	-0.007	-0.000
	(0.01)	(0.15)	(0.08)	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)
Married	-0.004	0.007	0.040	-0.000	0.001	0.004	-0.003	-0.000
	(0.00)	(0.04)	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Native	0.005	0.021	-0.011	-0.021	-0.006	-0.002	-0.015	0.000
	(0.01)	(0.20)	(0.12)	(0.04)	(0.01)	(0.01)	(0.01)	(0.00)
Father Employed	-0.022	0.650	0.268	-0.026	-0.020	0.005	-0.026	-0.002
	(0.01)	(0.35)	(0.20)	(0.04)	(0.01)	(0.02)	(0.01)	(0.00)
Twin	0.585	-3.674	-1.513	0.218	0.128	0.058	0.547	0.003
	(0.02)	(0.18)	(0.10)	(0.02)	(0.02)	(0.01)	(0.03)	(0.00)
Past Miscarriage	0.001	0.119	0.056	-0.003	0.001	0.006	0.002	-0.001
	(0.00)	(0.07)	(0.04)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Past Abortion	0.000	0.114	-0.037	-0.012	-0.010	0.006	0.003	0.000
	(0.00)	(0.08)	(0.05)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Metropolitan Area	0.003	-0.042	-0.092	-0.023	0.000	0.002	0.000	-0.000
	(0.00)	(0.08)	(0.04)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.025	50.847	34.855	0.234	0.095	0.055	0.044	0.009
	(0.04)	(0.86)	(0.50)	(0.17)	(0.05)	(0.04)	(0.04)	(0.01)
F-Test excluded instrument [*p*-value]	44.73 [0.00]	39.19 [0.00]	37.58 [0.00]	41.64 [0.00]	41.43 [0.00]	44.71 [0.00]	44.71 [0.00]	44.71 [0.00]
OLS results:								
Screening Test	-0.005	0.019	-0.039	-0.018	-0.012	-0.004	-0.006	-0.001
	(0.00)	(0.02)	(0.02)	(0.01)	(0.003)	(0.002)	(0.003)	(0.001)
Adj. R^2	0.11	0.06	0.02	0.01	0.01	0.004	0.10	0.001
N. Obs.	33,048	32,333	32,427	32,118	32,135	33,049	33,049	33,049

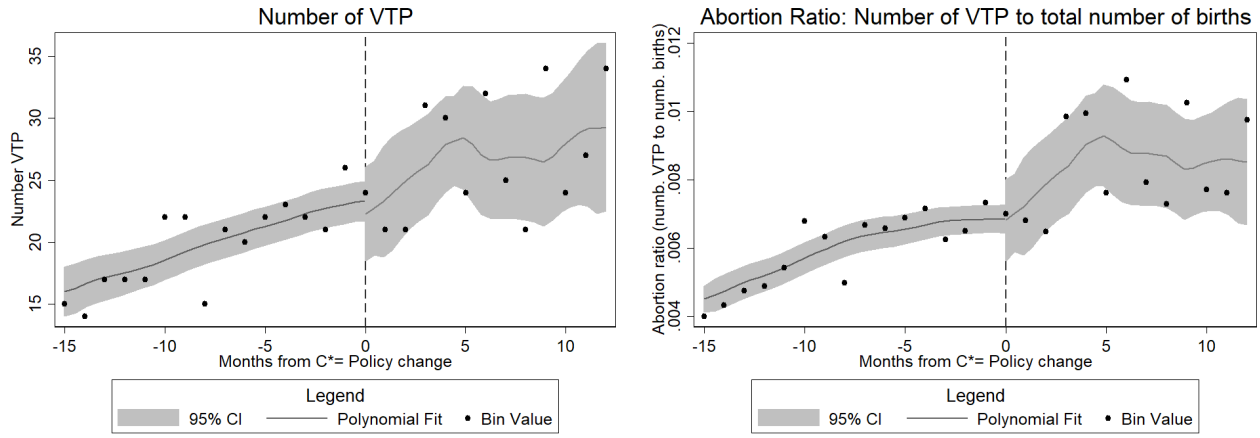
Notes: Each column presents the estimation of equation (2) by 2SLS. The *Screening Test* variable is a binary variable equal to one if a woman had a noninvasive screening test during pregnancy, and zero otherwise. *Assignment* A_i, equal to one in the post-policy period, is the instrumental variable for *Screening Test*. Standard errors (in round parentheses) are clustered at the level of district of residence–quarter of conception. OLS results for equation (2) are reported (Screening Test coefficient and Adj. R^2 only). Below the 'Screening test' coefficient, the table reports the original p-values (square brackets) and the "adjusted" p-values (in braces, the sharpened False Discovery Rate, FDR, q-values by Anderson, 2008), to account for multiple testing hypothesis.

Figure 1. Prenatal Tests Take-Up Rates Around the Policy Change



Notes: The y-axis measures the proportion of women undergoing no prenatal tests (top-left panel), a screening test (top-right panel), a diagnostic test (bottom-left panel), and both tests (bottom-right panel). The x-axis measures the number of weeks to (from) the policy change date (the zero value). Each dot represents the proportion for that week (Bin Average). The solid lines are the fit of a nonparametric polynomial regression model, separately estimated on both sides of the cutoff point (Polynomial Fit). The shaded grey area represents the 95 percent confidence interval (95 percent CI).

Figure 2. Voluntary Termination of Pregnancy Around the Policy Change



Notes: The y-axis measures the number of voluntary terminations of pregnancy (VTP, left panel), and the abortion ratio equal to the number of VTP to total number of births (right panel). The x-axis measures the number of months to (from) the policy change date (the zero value, October 2009). Each dot represents the number of VTP or the VTP ratio for that month (Bin Value). The solid lines are the fit of a nonparametric polynomial regression model, separately estimated on both sides of the cutoff point (Polynomial Fit). The shaded grey area represents the 95 percent confidence interval (95 percent CI). Data include all VTP that occurred during the second trimester of pregnancy, within the administrative borders of Piedmont, between July 2008 and October 2010. The source of data is the Italian Statistical Office (Istat).

Appendix

Table A1. Definition of Mothers' Characteristics

Variable	*Definition*
Age at Conception	Age at conception, expressed in years
Age group 18–24	Binary variable equal to one if age at conception is in the 18–24 range
Age group 25–29	Binary variable equal to one if the age at conception is in the 25–29 range
Age group 30–34	Binary variable equal to one if the age at conception is in the 30–34 range
Age group 35+	Binary variable equal to one if the age at conception is equal to or above 35
Low Education	Binary variable equal to one if the woman completed compulsory school or has no education
Medium Education	Binary variable equal to one if the woman completed high school
High education	Binary variable equal to if the woman has a university or higher degree
Employed	Binary variable equal to one if the woman is employed
Married	Binary variable equal to one if the woman is married
Native	Binary variable equal to one if the woman is born in Italy
Father Employed	Binary variable equal to one if the father is employed
Twin	Binary variable equal to one for a twin pregnancy
Past Miscarriage	Binary variable equal to one if the woman had past miscarriages
Past Abortion	Binary variable equal to one if the woman had past abortions
Metropolitan Area	Binary variable equal to one if the woman lives in a metropolitan area; Metropolitan areas are defined by the metropolitan area of Torino—the regional capital—and of the other seven largest towns of the region (provincial capitals).

Table A2. Summary Statistics of Mothers' Characteristics

	Whole Sample		**Pre-Policy**		**Post-Policy**	
	Mean	***SD***	***Mean***	***SD***	***Mean***	***SD***
Age at Conception	30.007	5.376	29.998	5.343	30.015	5.412
Age group 18–24	0.165	0.371	0.165	0.371	0.165	0.371
Age group 25–29	0.283	0.450	0.278	0.448	0.287	0.452
Age group 30–34	0.341	0.474	0.350	0.477	0.331	0.471
Age group 35+	0.211	0.408	0.206	0.404	0.216	0.412
Low Education	0.283	0.450	0.294	0.456	0.270	0.444
Medium Education	0.502	0.500	0.506	0.500	0.497	0.500
High education	0.216	0.411	0.200	0.400	0.233	0.423
Employed	0.711	0.453	0.707	0.455	0.716	0.451
Married	0.593	0.491	0.588	0.492	0.599	0.490
Native	0.766	0.423	0.771	0.420	0.760	0.427
Father Employed	0.910	0.286	0.898	0.302	0.923	0.267
Twin	0.010	0.101	0.009	0.094	0.012	0.108
Past Miscarriage	0.120	0.325	0.117	0.321	0.125	0.330
Past Abortion	0.075	0.264	0.071	0.256	0.080	0.271
Metropolitan Area	0.389	0.488	0.374	0.484	0.406	0.491

Notes: All variables are binary variables, except for Age at Conception, which is expressed in years.

Table A3. Test of the Smoothness of Mothers' Characteristics Around the Cutoff Date of the Policy

	Assignment A_i	*Std. Err.*
Dep. Var.		
Age at Conception	0.006	(0.02)
Age group 18–24	0.015	(0.02)
Age group 25–29	-0.015	(0.02)
Age group 30–34	-0.007	(0.02)
Age group 35+	-0.115	(0.28)
Low Education	0.007	(0.03)
Medium Education	-0.027	(0.03)
High education	0.020	(0.02)
Employed	-0.018	(0.03)
Married	0.004	(0.03)
Native	-0.005	(0.05)
Father Employed	-0.001	(0.03)
Twin	0.002	(0.003)
Past Miscarriage	-0.027	(0.01)
Past Abortion	-0.005	(0.01)
Metropolitan Area	0.009	(0.21)

Notes: In each row we estimate the following equation by OLS:

$$X_i = \tau_0 + \gamma A_i + \sum_{k=1}^{2} \tau_k (C_i - C^*)^k + \sum_{k=1}^{2} \pi_k A_i \times (C_i - C^*)^k + \mu_i$$

where each single observed mother characteristic X_i is considered a placebo outcome and is regressed on: 1) the treatment assignment dummy variable A_i equal to one in the post-policy period; 2) the term (C_i – C*), i.e. the number of weeks between the conception date C_i and the cutoff date C*, 3) the term (C_i – C*) squared, 4) the interaction of the trend (C_i – C*) with the treatment assignment variable. The table reports only the coefficient γ, for the post-policy dummy variable. Standard errors (in parentheses) are clustered at the level of district of residence–quarter of conception (138 clusters). Number of observations is 32,829 in all equations.

Table A4. Effect of Policy Change on Prenatal Tests' Take-Up Rates: Nonlinear Specifications

	(1)	(2)	(3)
	Quadratic	*Cubic*	*Quartic*
Dep. Var.: No Test			
Assignment A_i	-0.097	-0.107	-0.087
	(0.05)	(0.05)	(0.04)
Adj. R^2	0.09	0.09	0.09
Dep. Var.: Screening Test			
Assignment A_i	0.085	0.089	0.064
	(0.04)	(0.05)	(0.04)
Adj. R^2	0.10	0.10	0.10
Dep. Var.: Diagnostic Test			
Assignment A_i	0.012	0.012	0.010
	(0.01)	(0.01)	(0.02)
Adj. R^2	0.20	0.20	0.20
Dep. Var.: Both Tests			
Assignment A_i	0.0002	0.006	0.013
	(0.01)	(0.01)	(0.01)
Adj. R^2	0.02	0.02	0.02

Notes: Each row presents the estimation of equation (1) by OLS. The table reports only the post-policy dummy variable Assignment A_i. All equations include the trend ($C_i - C^*$), its squared term, its cubic term (only in columns 2 and 3), its quartic term (in column 4 only), and all the corresponding interactions with the treatment assignment variable. All equations also include the full set of observable characteristics for women (age, highest education level, mother's employment status, father's employment status, marital status, twin pregnancy, previous miscarriages and abortions, area of residence, and nationality). Standard errors (in parentheses) are clustered at the level of district of residence–quarter of conception (138 clusters). Number of observations is 32,829 in all equations.

Table A5. Effect of Policy Change on Prenatal Tests' Take-Up Rates: Alternative Bandwidths

	Bandwidth 26 weeks		***Bandwidth 104 weeks***	
	(1)	(2)	(3)	(4)
	Linear	*Quadratic*	*Linear*	*Quadratic*
Dep. Var.: No Test				
Assignment A_i	-0.096	-0.086	-0.077	-0.069
	(0.04)	(0.04)	(0.02)	(0.04)
Adj. R^2	0.08	0.08	0.10	0.10
Dep. Var.: Screening Test				
Assignment A_i	0.084	0.067	0.071	0.056
	(0.04)	(0.04)	(0.02)	(0.03)
Adj. R^2	0.10	0.10	0.11	0.11
Dep. Var.: Diagnostic Test				
Assignment A_i	0.012	0.008	0.006	0.012
	(0.01)	(0.01)	(0.01)	(0.01)
Adj. R^2	0.20	0.20	0.19	0.19
Dep. Var.: Both Test				
Assignment A_i	0.001	0.011	-0.0004	0.0004
	(0.01)	(0.01)	(0.004)	(0.01)
Adj. R^2	0.02	0.02	0.02	0.02
N. Obs.	17,228	17,228	64,703	64,703

Notes: Each row presents the estimation of equation (1) by OLS. The table reports only the post-policy dummy variable Assignment A_i. All equations include the trend ($C_i - C^*$), its squared term (only in columns 2 and 4), and all the corresponding interactions with the treatment assignment variable. All equations also include the full set of observable characteristics for women (age, highest education level, mother's employment status, father's employment status, marital status, twin pregnancy, previous miscarriages and abortions, area of residence, and nationality). Standard errors (in parentheses) are clustered at the level of district of residence-quarter of conception.

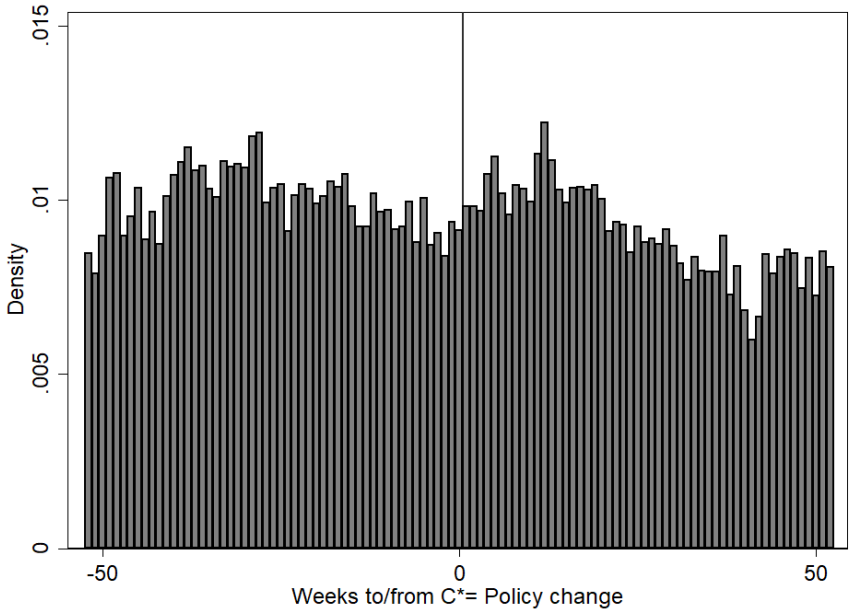
Table A6. Effect of Policy Change on the Prenatal Tests' Take-Up Rates: Nonparametric Specifications

	(1)	(2)	(3)	(4)
	Linear	*Quadratic*	*Cubic*	*Quartic*
Dep. Var.: No Test				
Assignment A_i	-0.092	-0.100	-0.099	-0.098
	(0.04)	(0.05)	(0.05)	(0.04)
bandwidth	39	44	53	54
Dep. Var.: Screening Test				
Assignment A_i	0.080	0.083	0.075	0.078
	(0.04)	(0.04)	(0.04)	(0.04)
bandwidth	40	44	48	54
Dep. Var.: Diagnostic Test				
Assignment A_i	0.010	0.011	0.012	0.012
	(0.01)	(0.01)	(0.01)	(0.02)
bandwidth	50	56	57	69
Dep. Var.: Both Tests				
Assignment A_i	0.0002	0.002	0.010	0.010
	(0.01)	(0.01)	(0.01)	(0.01)
bandwidth	46	54	47	72

Notes: Each row presents the estimation of equation (1). Estimation method is local linear and polynomial (up to the fourth order) nonparametric in all columns. The optimal bandwidth is chosen by the MSE-optimal bandwidth selector, and the observations are weighted by a triangular kernel. The table reports only the post-policy dummy variable Assignment A_i. All equations also include the full set of observable characteristics for women. Standard errors (in parentheses) are clustered at the level of district of residence–quarter of conception.

Figure A1. Distribution of Pregnant Women

Panel A. Histogram of pregnant women density according to the conception week, around the policy change.



Panel B. Histogram of pregnant women density according to the age at conception (vertical line is age 35).

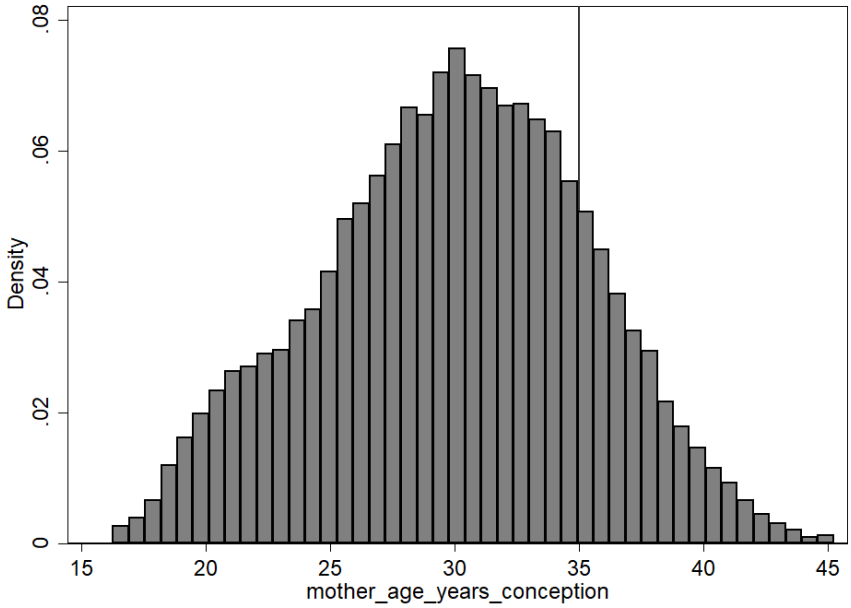


Figure A2. Mothers' Characteristics Around the Policy Change

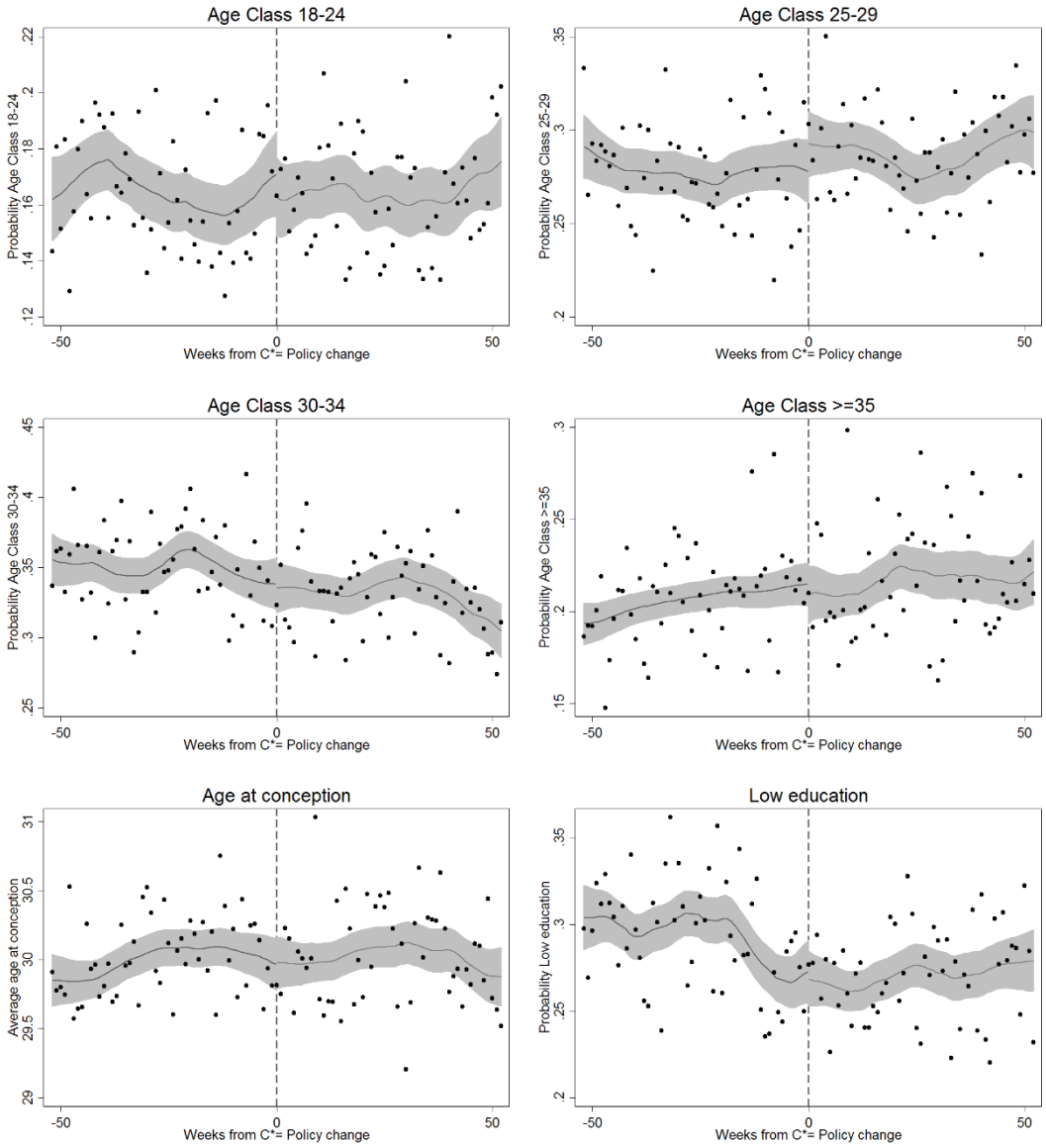
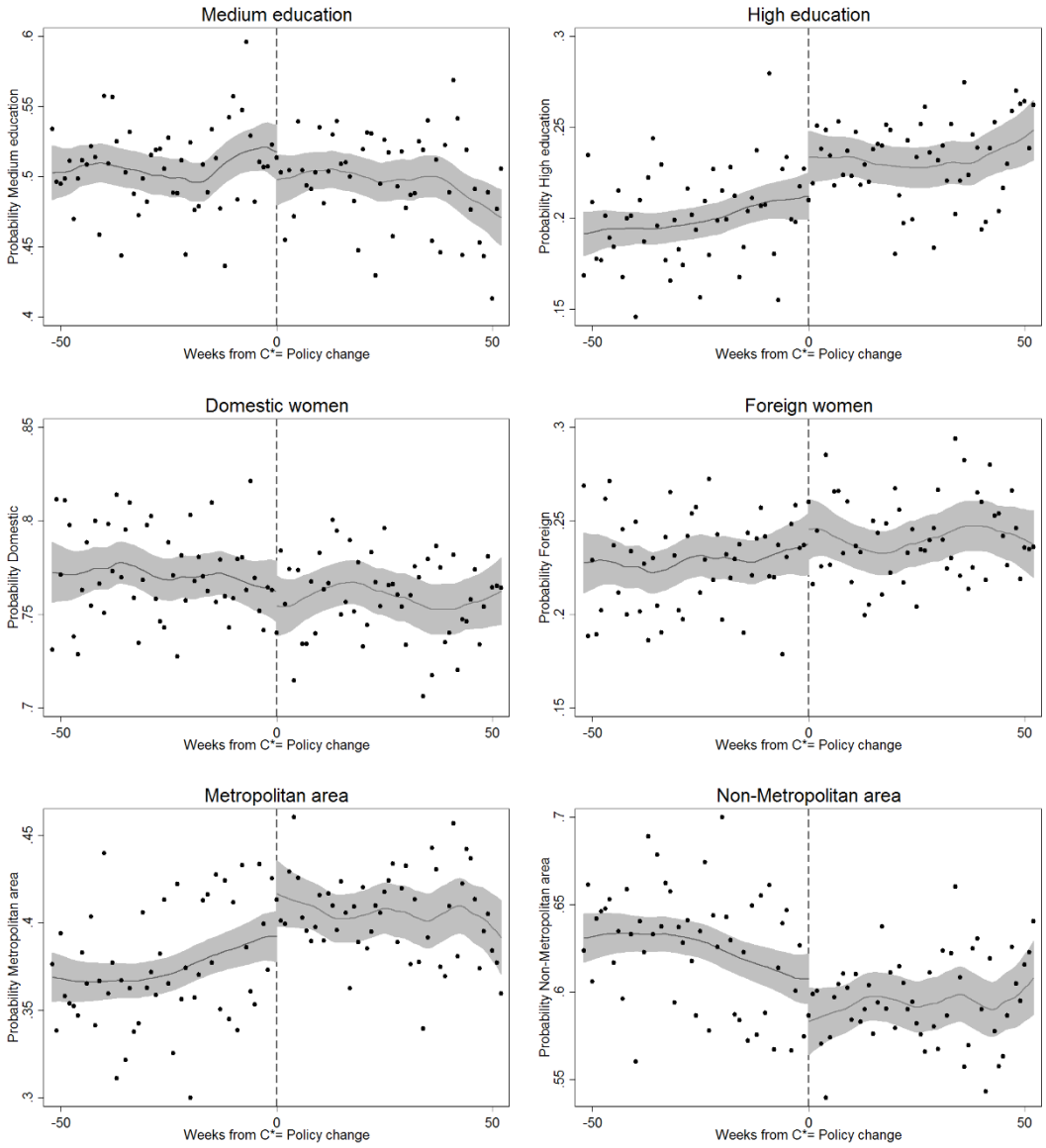
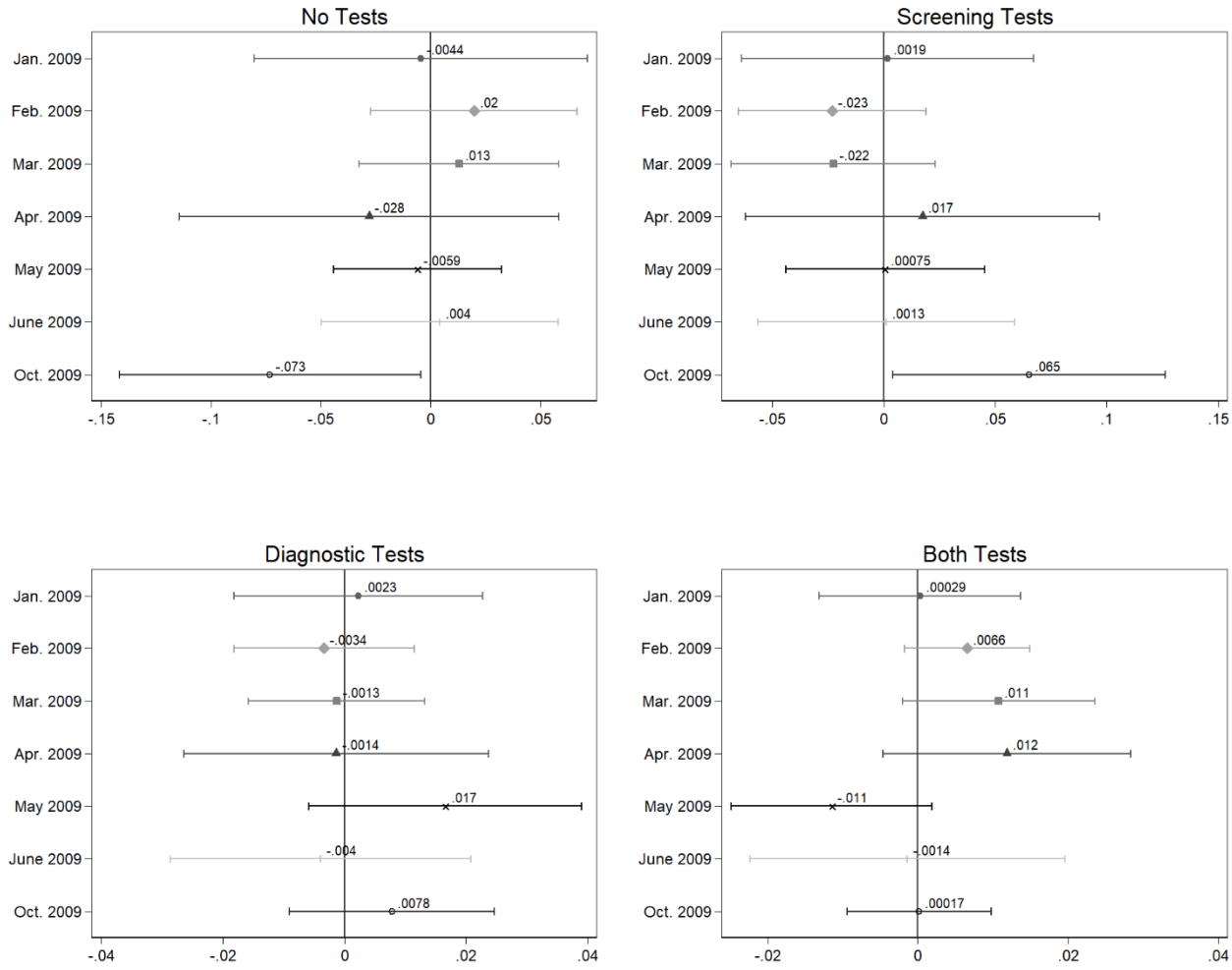


Figure A2. Mothers' Characteristics Around the Policy Change (continued)



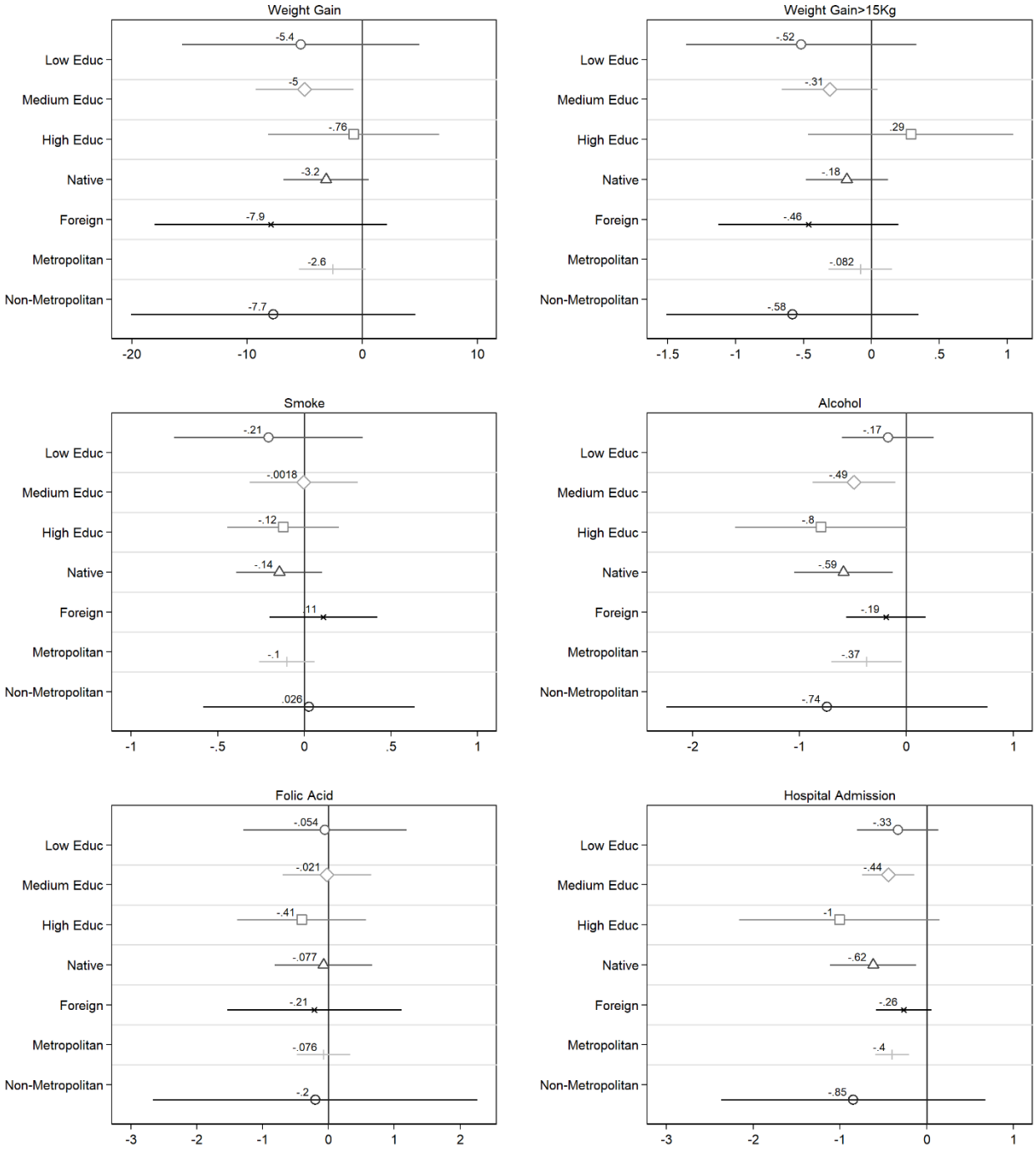
Notes: The y-axis measures the mothers' characteristics. The x-axis measures the number of weeks to (from) the policy change date (the zero value). Each dot represents the average mother's characteristics for that week (Bin Average). The solid lines are the fit of a nonparametric polynomial regression model, separately estimated on both sides of the cutoff point (Polynomial Fit). The shaded grey area represents the 95 percent confidence interval (95 percent CI).

Figure A3. Falsification Test: Estimated Treatment Effects for Cutoff Dates Artificially Anticipated



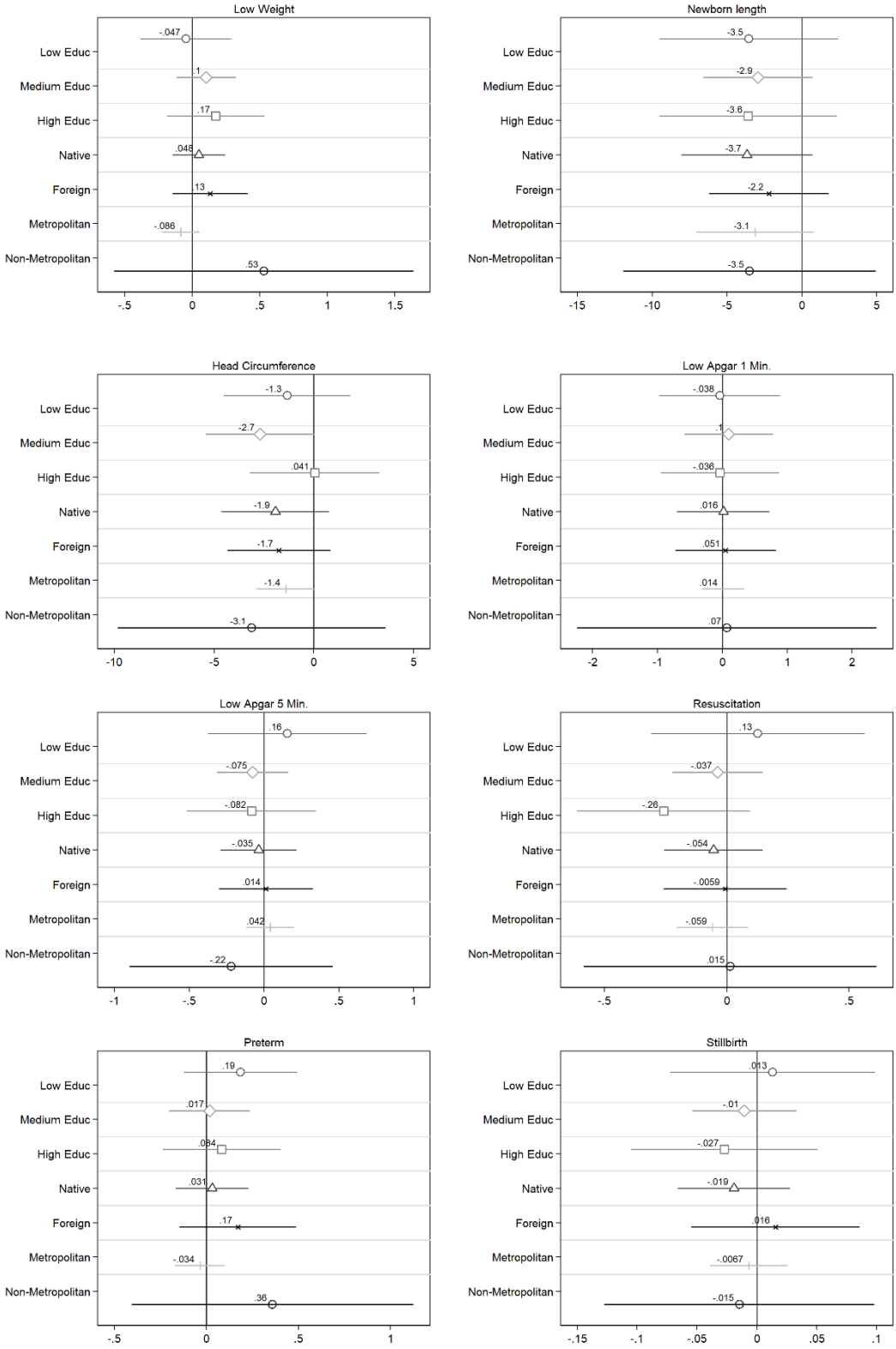
Notes: Each plot shows seven replications of equation (1). The dependent variables are no prenatal tests (top-left panel), screening test (top-right panel), diagnostic test (bottom-left panel), and both tests (bottom-right panel). In each replication, we consider a different cutoff date for the policy change, from January 1, 2009 to October 1, 2009, at four-week intervals. We exclude women who became pregnant in July, August, or September 2009. The plots show the point estimates and the 95 percent confidence intervals. The x-axis measures the estimated coefficients for the treatment assignment A_i, a dummy variable equal to one in the post-policy period, in equation (1), for the linear specification only. The y-axis reports the considered cutoff date. The estimated coefficients labeled Oct. 2009 is the true cutoff date.

Figure A4. The Effect of Screening Tests on Mother's Health behaviors by Subsample



Notes: Each plot shows 2SLS results of equation (2) for seven subsamples of women (low educated, medium educated, high educated, native, foreign born, living in metropolitan areas, and living in nonmetropolitan areas). The dependent variables are weight gain during pregnancy and weight gain during pregnancy larger than 15 kg (top panels), smoking during pregnancy and consuming alcohol during pregnancy (middle panels), and folic acid supplements during pregnancy and hospital admission in pregnancy (bottom panels). The plots show the point estimates and the 95 percent confidence intervals. The x-axis measures the estimated coefficients for the screening test, a dummy variable equal to one if the woman had a screening test, and zero otherwise. The instrumental variable for the screening test variable is the post-policy dummy variable Assignment A_i. The y-axis reports the considered subsamples.

Figure A5. The Effect of Screening Tests on Newborn Health Outcomes by Subsample



Notes: Each plot shows 2SLS results of equation (2) for seven subsamples of women (low educated, medium educated, high educated, native, foreign born, living in metropolitan areas, and living in nonmetropolitan areas). The dependent variables are health outcomes for newborns. The plots show the point estimates and the 95 percent confidence intervals. The x-axis measures the estimated coefficients for the screening test, a dummy variable equal to one if the woman had a screening test, and zero otherwise. The instrumental variable for the screening test variable is the post-policy dummy variable Assignment A_i. The y-axis reports the considered subsamples.