NBER WORKING PAPER SERIES

WHEN MORE PRENATAL CARE LOOKS LIKE LESS: SELECTION, INFORMATION AND MATERNAL AND INFANT HEALTH

Chien-Tzu Cheng
Shin-Yi Chou
Hsien-Ming Lien

Working Paper 36000
http://www.nber.org/papers/w36000

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2026

Cheng gratefully acknowledges financial support from PSTC International Fellowship through the Population Studies and Training Center at Brown University, which receives funding from the NIH for training support (T32 HD007338) and for general support (P2C HD041020) and from MEXT/ JSPS KAKENHI Grant Number 25K16665. Lien is grateful for financial support from Taiwan National Science and Technology Council (Grant #: 113-2410-H-004-107-MY2). We are thankful for the administrative data provided by the National Health Insurance Agency. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Chien-Tzu Cheng, Shin-Yi Chou, and Hsien-Ming Lien. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

When More Prenatal Care Looks Like Less: Selection, Information and Maternal and Infant Health

Chien-Tzu Cheng, Shin-Yi Chou, and Hsien-Ming Lien

NBER Working Paper No. 36000
November 2026
JEL No. I12, I18, J13

ABSTRACT

We use Taiwan's 2003 SARS epidemic as a natural experiment, instrumenting prenatal care with gestational exposure to SARS-induced fear that reduced visits by over 20 percent. Adequate prenatal care reduces preventable maternal complications by 3.4 percentage points (26 percent), with no effect on infant outcomes. OLS estimates are perversely signed for maternal complications but overstate benefits for infant outcomes, consistent with competing patient- and provider-driven selection. Mechanism evidence points to prenatal care as a longitudinal information system: repeated contact facilitates detection of evolving risks and generates clinical information that informs delivery decisions, including delivery-mode routing and subsequent management at delivery.

Chien-Tzu Cheng
The University of Osaka
Graduate School of Economics
chien-tzu_cheng@econ.osaka-u.ac.jp

Hsien-Ming Lien
National Cheng-Chi University
Department of Public Finance
hmlien@nccu.edu.tw

Shin-Yi Chou
Lehigh University
Department of Economics
and NBER
syc2@lehigh.edu

1 Introduction

Prenatal care is one of the most widely used interventions for improving maternal and infant health, providing opportunities for early risk detection, nutritional guidance, linkage to safe delivery, and prevention of avoidable complications (World Health Organization, 2016; Campbell and Graham, 2006). Yet despite its widespread use, evidence on its causal effects remains mixed. Studies using quasi-experimental and within-family research designs generally find beneficial effects,[1] but estimated magnitudes vary substantially across settings, outcomes, and research designs, with some studies finding little or no effect.[2] A key challenge in interpreting these estimates is that selection into prenatal care can operate in opposing directions. Women who are more health-conscious may seek more care and have better outcomes for reasons unrelated to prenatal care, leading observational estimates to overstate its benefits. Conversely, providers may direct higher-risk pregnancies toward more intensive monitoring, leading observational estimates to understate the benefits of care.

These opposing selection forces complicate both the estimation and interpretation of prenatal care effects: observational estimates may differ from causal effects, and the direction of the bias may itself differ across outcomes. This motivates two related questions. First, what is the causal effect of adequate prenatal care on maternal and infant health? Second, how does selection into prenatal care differ across outcomes?

We address these questions using Taiwan's 2003 SARS epidemic as a natural experiment. The outbreak generated widespread fear of healthcare settings and a sharp, unanticipated decline in prenatal care utilization. Average prenatal visits fell from 9.3 to 7.3, despite Taiwan's National Health Insurance program covering ten prenatal visits during this period. Because pregnancies differed in how much of gestation overlapped with the outbreak window, the timing of SARS generated plausibly exogenous differences in the disruption to prenatal care across pregnancies. Using linked administrative records covering approximately 335,000 births in 2002–2003, we instrument prenatal care adequacy, measured by the Kessner index, with the share of gestation overlapping the SARS period.

Our results reveal a striking asymmetry between maternal and infant outcomes. Adequate prenatal care reduces preventable maternal complications at delivery by 3.4 percentage points, or 26

[1] These approaches include instrumental variables strategies exploiting exogenous shocks to care access, such as bus strikes (Evans and Lien, 2005), Medicaid expansions (Currie and Grogger, 2002), weather shocks (Gajate-Garrido, 2013), and marital status (Jewell and Triunfo, 2006); quasi-experimental designs leveraging discrete policy changes such as Medicaid reimbursement rate changes (Sonchak, 2015), cash incentives for early care initiation (Cygan-Rehm and Karbownik, 2022), and improvements in neighborhood safety (Lautharte, 2021); and mother fixed-effects models that compare siblings within families to control for time-invariant maternal characteristics (Abrevaya and Dahl, 2008; Yan, 2017, 2020).

[2] Corman et al. (2019) documents several systematic patterns in this literature. Estimated effects tend to shrink as identification strategies become more rigorous, consistent with declining positive selection in more recent cohorts with greater access to prenatal care; effects tend to be larger in developing-country or disadvantaged settings where baseline access is lower (e.g., Gajate-Garrido, 2013; Cygan-Rehm and Karbownik, 2022; Lautharte, 2021; Currie and Grogger, 2002); and mother fixed-effects estimates tend to exceed instrumental-variables estimates, potentially because within-mother variation in prenatal care may partly reflect responses to time-varying health shocks.

percent, but has no detectable effect on small-for-gestational-age birth, low birth weight, or neonatal and infant mortality. More importantly, observational and instrumental-variable estimates diverge in opposite directions across these outcomes. OLS suggests that women receiving more prenatal care experience more maternal complications, masking the protective causal effect that emerges under IV. For infant outcomes, by contrast, OLS suggests substantial benefits that disappear under IV. Thus, within the same population and empirical setting, selection leads observational estimates to understate the benefits of prenatal care for maternal complications while overstating them for infant outcomes. A series of negative-control, subgroup, and regional placebo tests provides no evidence that direct effects of SARS, rather than disrupted prenatal care, account for these findings.

We interpret this pattern as evidence that two distinct forms of selection may coexist. Patient-driven selection arises when women differ in their propensity to seek prenatal care for reasons that are also related to health outcomes, whereas provider-driven selection arises when clinicians respond to observed or emerging pregnancy risks by recommending more intensive prenatal monitoring. In our setting, health-conscious care seeking provides one plausible form of patient-driven selection and is consistent with the favorable OLS associations for infant outcomes. Provider responses to maternal risk, by contrast, are consistent with the unfavorable OLS association for maternal complications. Provider-driven selection may be particularly relevant for maternal outcomes because many risk factors for complications at delivery are observable to providers at baseline or become apparent through prenatal monitoring. Providers may respond to these risks with additional visits or closer monitoring, creating a positive association between prenatal care utilization and underlying maternal risk even when additional care is protective. The direction of observational bias may therefore depend on the relative importance of patient-driven care seeking and provider responses to clinical risk.

The role of clinical information in provider-driven selection also points to a potential mechanism through which prenatal care itself improves maternal outcomes: the accumulation of information over the course of pregnancy. Our mechanism results suggest that an important function of prenatal care is to generate longitudinal clinical information, rather than simply to deliver a sequence of discrete medical services. Repeated prenatal encounters give providers opportunities to detect evolving conditions and incorporate that information into delivery decisions. Consistent with this interpretation, adequate prenatal care increases the recorded diagnosis of anemia, which can develop after the routine first-trimester blood screen and may therefore require subsequent testing to detect, while having little effect on diagnoses of conditions more readily identified through routine monitoring. Adequate care also increases cesarean delivery by 7.2 percentage points, suggesting that information accumulated during pregnancy changes how delivery is managed. A mechanical decomposition indicates that this shift in delivery mode can account for roughly 40 percent of the reduction in preventable maternal complications. At the same time, prenatal care also reduces complications among vaginal deliveries, indicating that its benefits extend beyond the choice of delivery mode.

This paper contributes to three literatures. First, we contribute to the quasi-experimental

literature on prenatal care, whose estimated effects vary widely in magnitude and occasionally diverge in sign (Corman et al., 2019). Rather than adding another point estimate to this literature, we show that this heterogeneity need not arise solely from differences in methods, settings, or baseline access: the direction of selection itself may differ across outcomes. Because we examine maternal and infant outcomes within the same sample and identification strategy, we can observe opposing patterns of selection that studies focusing on a single class of outcomes may miss. More broadly, our findings illustrate how differences between OLS and IV estimates can be informative about the nature of selection into treatment (Bjerk, 2025; Ciacci, 2025; Black et al., 2022).

Second, we contribute more broadly to the literature on selection into medical care. Existing work emphasizes how clinical recommendations interact with patient behavior to shape selection into care (Einav et al., 2020; Oster, 2020). We highlight a complementary dimension: when treatment intensity is jointly determined by patients' care-seeking behavior and providers' responses to emerging clinical information, the direction of selection can vary across outcomes within the same episode of care. This interaction between patient- and provider-driven selection may extend beyond prenatal care to other settings involving repeated medical contact and evolving information about patient risk.

Third, we contribute to understanding the informational value of repeated healthcare encounters. Our setting builds most closely on Bennett et al. (2015), who document SARS-induced declines in aggregate outpatient utilization in Taiwan, as well as the COVID-era literature on prenatal stress and care disruption (Liu et al., 2020, 2022). Rather than treating the epidemic simply as a shock to the quantity of healthcare consumed, we exploit the structured and clinically timed nature of prenatal care to study what is lost when routine contact is disrupted. We show that SARS disrupted prenatal care throughout pregnancy but disproportionately reduced care in the second and third trimesters, including a sharp decline in use of the scheduled second-trimester anomaly scan. Because prenatal care provides different forms of screening and surveillance at different stages of pregnancy, these patterns allow us to connect disruptions in routine contact to losses of clinical information and, ultimately, to decisions made at delivery. More broadly, our findings highlight an important value of repeated healthcare encounters beyond the services delivered at each visit: the longitudinal information they generate for subsequent clinical decision-making.

The remainder of the paper proceeds as follows. Section 2 describes the SARS outbreak and its effects on healthcare utilization in Taiwan. Section 3 describes our data, sample, and key variables. Section 4 presents our empirical strategy. Section 5 reports first-stage, reduced-form, and instrumental variables estimates. Section 6 addresses threats to the exclusion restriction. Section 7 presents evidence on mechanism. Section 8 concludes.

2 Background

2.1 The SARS Epidemic in Taiwan

SARS, the first severe and readily transmissible disease of the 21st century, emerged in the Guangdong province of China on November 16, 2002, and spread internationally through travel networks.[3] By the time the WHO declared global containment on July 5, 2003, the virus had caused approximately 800 deaths across more than 30 countries and an estimated US$40 billion in economic losses (World Health Organization, 2003; Lee et al., 2004; Oberholtzer et al., 2004).

Figure 1 summarizes the key epidemic milestones in Taiwan. The country's first confirmed case was reported on March 14, 2003—the date we adopt as the start of the SARS period. The outbreak escalated sharply on April 22, when a nosocomial cluster at Heping Hospital in Taipei, eventually linked to 26 deaths, triggered facility quarantine and a nationwide public health emergency. By May 3, cumulative cases had surpassed 100, prompting the WHO to dispatch investigative teams; on May 21, the WHO issued a travel advisory for Taiwan, amplifying public anxiety well beyond the directly affected areas. As of May 22, Taiwan had reported 483 probable cases and 60 deaths concentrated overwhelmingly in Taipei City and Taipei County (Appendix Figure A1.1), making it the world's third-largest outbreak after China and Hong Kong.[4] The final toll stood at 674 probable cases and 87 deaths.[5] Notably, fear persisted well beyond formal containment: prenatal care utilization did not recover to pre-epidemic levels until after 2006 (Appendix Figure A1.2).

2.2 Fear of SARS and the Disruption of Prenatal Care

Beyond its direct mortality, SARS generated a pervasive fear of healthcare settings that sharply curtailed medical utilization. This fear was disproportionate to individual risk: although the disease killed some 13 percent of those it infected, with no vaccine, treatment, or rapid test available, the underlying incidence was minute—at roughly three cases per 100,000, almost no one in Taiwan encountered SARS directly. Yet what transmission did occur was concentrated in hospitals, as the Heping cluster made vivid, so the institutions people relied on for care were the sites where risk was highest. As during the later COVID-19 pandemic, school and hospital closures, mass quarantines, and public health campaigns warning against unnecessary outings led patients to defer even routine care (Maunder, 2004; Nickell et al., 2004). Bennett et al. (2015) documents that fear of SARS reduced total outpatient visits in Taiwan by 30 percent.

Prenatal care was no exception. Our study period spans 2002-2003, and Figure 2a plots monthly prenatal visit rates over this window. We define births on or after March 14, 2003 as the treated group, and the treatment period as March 14 to December 31, 2003. The average number of prenatal visits per birth fell from 9.3 in the pre-treatment period to 7.3 in the treatment period, a

[3] The Chinese government did not officially notify the WHO of the outbreak until February 10, 2003, nearly three months after the first identified case. See here.

[4] Revised WHO and Taiwan CDC estimates place lab-confirmed cases at 346, with 279 in the Taipei metropolitan area. See here.

[5] See here.

decline of more than 20 percent of the ten visits covered by national health insurance. Crucially, the drop preceded the large outbreak in late April, coinciding instead with initial news coverage of confirmed cases—consistent with a response to information rather than exposure. As case counts rose, visit rates fell further, remaining depressed through the end of 2003, well after the WHO declared containment. Because this decline was unanticipated by expectant mothers and unrelated to their underlying health, it provides a plausible source of exogenous variation in prenatal care utilization.

2.3 Institutional Setting: National Health Insurance and Prenatal Care

Taiwan's National Health Insurance (NHI), established in 1995, provides near-universal single-payer coverage to the country's 23 million residents, including outpatient visits, inpatient care, prescription drugs, and dental services. Cost-sharing is minimal—outpatient co-payments are approximately US$5 at clinics and US$8 at hospitals—and patients may seek care at any NHI-affiliated provider.[6] The near-zero out of pocket expense means that financial barriers to prenatal visits are negligible, so reductions in utilization during the SARS period are unlikely to reflect cost constraints.

NHI covers ten prenatal care visits to expectant mothers (see Figure A2.1). The schedule calls for two visits in the first trimester (weeks 0–16), two in the second (weeks 17–28), and six in the third (weeks 29–40). Each visit follows a standardized protocol recorded in the "Mother's Book," a government-issued booklet distributed at enrollment that serves as the official longitudinal health record for the pregnancy.

Three visits carry particular clinical significance: Visit 1 (prior to week 12) establishes a baseline and screens for early infections; Visit 3 (week 20) performs the second-trimester anomaly scan (or detailed anomaly ultrasound) for fetal anomalies; and Visit 5 (weeks 29–32) reassesses infection risk (e.g. Hepatitis B antibody test; HBsAG, HBeAG, and VDRL laboratory tests), fetal position, and gestational diabetes. A complete description of visit content across all ten check-ups and their clinical rationale is provided in Appendix A2.

Because the schedule is structured and government-mandated, a missed visit corresponds to a specific, identifiable set of foregone clinical services rather than a generic reduction in care. This link is what gives the SARS-induced decline in utilization a well-defined clinical mechanism: pregnant women exposed to the epidemic during particular gestational windows would have skipped the visits carrying diagnostic information unique to that window, rather than an interchangeable share of overall care.

[6] Inpatient cost-sharing is capped at 10 percent of the average national income per capita. Certain groups, including indigenous peoples (roughly 2 percent of the population), are exempt from all co-payments.

3 Data, Sample and Variables

3.1 Data and Sample

Our analysis draws on three administrative data sources from Taiwan covering 2001–2003. The primary source is the full-population longitudinal claims data from the National Health Insurance Research Database (NHIRD), which contains all inpatient and outpatient claims submitted to NHI system. The database includes individual identifiers, ICD-9 diagnosis codes, diagnosis-related group (DRG) codes, hospital identifiers, dates of admission and service, and patient characteristics such as birth dates and place of residence. We supplement the claims data with two additional administrative sources. Birth certificate records provide information on birth weight, gestational age—calculated from the date of the last menstrual period and reported in completed weeks—and maternal education. Death certificate records are linked to birth records using individual identifiers to measure neonatal and infant mortality. We include claims from 2001 to capture prenatal visits occurring before delivery among women who gave birth early in 2002, since these pregnancies began several months before the start of our delivery window.

We construct the study sample by identifying all women who delivered between 2002 and 2003 and linking their delivery records to the corresponding birth certificates.[7] This two-year delivery window covers the period immediately surrounding the 2003 SARS outbreak. We then use mothers' outpatient claims to identify prenatal care, both free and paid visits, received during pregnancy.[8]

We impose several sample restrictions. First, non-citizen mothers or mothers with fewer than three NHI-covered prenatal visits are excluded because their prenatal care records are likely incomplete, possibly due to not residing in Taiwan throughout the full gestational period. Next, to ensure comparability of health outcomes across the sample, we exclude multiple births, parity above six, birth weights below 400 or above 6,500 grams, maternal age outside 15–40 years, and gestational lengths outside 29–42 weeks. Deliveries on offshore islands—Kinmen, Penghu, and Matsu—are also excluded due to limited and unrepresentative healthcare access; these account for fewer than 1% of observations. Finally, we drop observations with missing maternal education or ZIP code identifiers. These restrictions yield a final analysis sample of 335,674 births.

3.2 Variables

Outcome Variables.—We examine two sets of outcomes: preventable maternal delivery complications and infant birth outcomes. Maternal outcomes are constructed from the NHIRD claims data, while infant outcomes are drawn from birth and death certificates.

For maternal outcomes, we construct a binary indicator for experiencing any preventable com-

[7] Vaginal and cesarean deliveries are identified from inpatient claims using DRG codes 0371A, 0373A, 0373B, and 0373C.

[8] Prenatal care visits are identified using NHI prenatal care codes IC40–IC50 and the ICD-9 code for normal pregnancy supervision (V22). This definition captures both NHI-covered prenatal visits and additional privately paid prenatal visits recorded in the claims data.

plication during labor, combining three categories: infection, injury, and bleeding.[9] We focus on conditions that adequate prenatal care can prevent or mitigate: infections that are detectable and treatable through prenatal screening, such as urinary tract infections (Silverman and Turrentine, 2023) and syphilis (US Preventive Services Task Force, 2018); birth canal injuries and maternal trauma that can be mitigated by identifying high-risk cases, such as prior cesarean deliveries, before labor (Landon et al., 2005); and antepartum hemorrhage not attributable to an identified placental abnormality, the subset of bleeding without a fixed anatomical cause and therefore the subset where antenatal management is most plausibly consequential. We deliberately exclude hemorrhage arising from identified placental causes—placenta previa, abruption, and coagulation defects—whose occurrence is driven largely by fixed features of the pregnancy rather than by conditions modifiable through prenatal care.[10] Because each category is individually rare, we combine them into a single composite indicator to improve statistical precision.

For infant outcomes, we examine four measures constructed from birth and death certificates: small for gestational age (SGA), low birth weight (LBW), neonatal mortality, and infant mortality. SGA is defined as a birth weight below the 10th percentile for gestational age by week, and LBW as a birth weight below 2,500 grams, regardless of gestational age. While both measures are commonly used to capture fetal growth restriction, they reflect partially distinct dimensions: LBW conflates growth restriction with prematurity, since it can result from either a shortened gestation or restricted growth during a normal-length pregnancy, whereas SGA, by conditioning on gestational age, more directly isolates the growth dimension, though it also captures some infants who are constitutionally small rather than growth-restricted. Both outcomes are nonetheless clinically relevant to the adequacy of prenatal care, which can include nutrition supplementation, management of conditions such as maternal anemia, and counseling on smoking and other substance use during pregnancy, interventions with documented effects on birth weight and growth outcomes. Neonatal mortality denotes death within the first 28 days of life and infant mortality denotes death within the first year; adequate prenatal care is similarly thought to contribute to both by identifying and managing complications that elevate the risk of early infant death, though whether this translates into detectable average effects in a general population remains an open empirical question, one we return in our results.

Prenatal Care Measures.—Our primary treatment variable is the adequacy of prenatal care, measured using the Kessner Adequacy of Prenatal Care Index (Kessner et al., 1973). Originally developed to evaluate the adequacy of prenatal care utilization in population-level studies, the Kessner index has been widely adopted in the health economics and public health literature as a summary measure of prenatal care quality (Corman et al., 2019). The index jointly incorporates

[9] These categories capture a broader range of ICD-9-coded conditions than the illustrative examples given here; see Appendix Table A6.1 (Panel A) for the ICD-9 reference codes used to identify obstetric infection, injury, and antepartum hemorrhage.

[10] This exclusion is conservative in one respect: detection of placenta previa on second-trimester ultrasound permits planned cesarean delivery, so prenatal care may avert hemorrhage even when it cannot prevent the underlying condition. Excluding these cases therefore understates the scope for prevention. We prefer this to the alternative, since including them would attribute to prenatal care variation driven by placental pathology it cannot influence.

three dimensions of care utilization: the trimester in which care was initiated, the total number of prenatal visits received, and gestational age at delivery. By combining initiation timing with visit counts conditional on gestational length, the index accounts for the fact that a given number of visits carries different implications depending on how early care began and how long the pregnancy lasted. Based on these criteria, the index classifies each woman's prenatal care utilization into one of three categories: adequate, intermediate, or inadequate. We define our treatment indicator as equal to one if a woman's care is classified as adequate and zero otherwise, pooling intermediate and inadequate utilization as the comparison group.

Figure 2b illustrates the evolution of the Kessner adequacy index (left axis) over time relative to the timing of birth. Prior to the onset of SARS, approximately 70% to 80% of pregnant women received adequate prenatal care; this share declined steadily since the outbreak hit, reaching about 40% by December 2003. The pattern observed using the Kessner index closely mirrors the trend in the raw number of prenatal visits shown in Figure 2a. Taken together, these patterns imply that one major source of variation in prenatal visits was predominantly induced by SARS. For robustness, we employ two alternative measures—an indicator for receiving at least 10 visits (capturing compliance with the recommended threshold) and the total number of visits (capturing the intensity of care use)—with results shown in Appendix A4.

3.3 Exposure to SARS

To address the endogeneity of prenatal care utilization, we exploit the SARS outbreak as an exogenous shock to care access. We construct a continuous instrument equal to the fraction of each woman's pregnancy that overlapped with the period we define as the SARS window, beginning with Taiwan's first confirmed case on March 14, 2003, and ending December 31, 2003, the end of our sample period. Because the broader epidemic's effects on healthcare-seeking behavior persisted well beyond any formal epidemiological endpoint (Appendix Figure A1.2), we do not attempt to define an objective outbreak duration and instead anchor the window to dates we can observe directly in our sample:

$$\text{Exposure fraction (IV)} = \frac{\text{Number of exposed days during pregnancy}}{\text{Length of SARS window (292 days)}}. \tag{1}$$

The measure takes a value of 0 for mothers who conceived and delivered entirely before the outbreak, and 1 for mothers who conceived before the outbreak but delivered after it ended, such that their pregnancy overlapped with the SARS period throughout. Figure 2b displays the joint evolution of the Kessner adequacy index (left axis) and the SARS exposure fraction over birth timing (right axis), illustrating the central identifying pattern: as the exposure fraction rises monotonically with birth timing after March 2003, the Kessner adequacy index moves in the opposite direction. This inverse relationship between gestational SARS exposure and prenatal care adequacy forms the basis of our first stage and confirms that the outbreak generated substantial and systematic variation in care utilization.

Approximately 97% of our sample conceived before the outbreak's start date, so the intensity of each woman's gestational exposure to SARS was determined by her delivery date relative to the outbreak rather than by any deliberate choice to conceive. Two potential selection issues remain, however, which we address in turn.

The first involves the remaining 3% of our sample who conceived after March 14, 2003. Dropping this group, as one might be tempted to do, would introduce a more serious problem than it solves. Our sampling frame is a birth cohort: a pregnancy enters the sample if and only if it results in a delivery recorded on or before December 31, 2003, and every such delivery is observed regardless of gestational length. Conception timing is therefore not independently observed but a deterministic function of two other variables, $C_i = B_i - G_i$, where B_i is the date of delivery and G_i gestational age at birth, dated from the last menstrual period. Restricting the sample to $C_i <$ March 14, 2003 is thus equivalent to imposing $G_i \geq B_i -$ March 14, 2003—a restriction on gestational age, that is, on a birth outcome. Because we already exclude deliveries before 29 weeks, this screen is vacuous through October 3, 2003 and binds only thereafter, with a threshold rising one-for-one with the date of delivery: 33 weeks for deliveries on November 1, 37 weeks on December 1, and 42 weeks by December 31. It would therefore leave the weakly exposed cohorts untouched while stripping the short-gestation tail from deliveries in the final quarter of 2003, precisely the cohorts carrying the highest gestational SARS exposure. The surviving high-exposure cells would consist disproportionately of pregnancies that reached term, attenuating the reduced form and biasing our estimates of the effect of prenatal care toward zero. Conditioning on conception date is conditioning on a post-treatment variable, and we therefore retain the full birth cohort.[11] Two features of the setting further reassure us that doing so does not reintroduce endogenous fertility timing: the outbreak was sudden and unanticipated, leaving little time to adjust medium-term plans such as conception decisions, and conception is in any case subject to considerable biological uncertainty.

The second issue is of a different nature and involves women absent from our sample altogether: those who deliberately postponed pregnancy and conceived only after the epidemic subsided. This concerns who becomes pregnant rather than which pregnancies we observe, and it bears on the exclusion restriction rather than on sample selection. To the extent that the women who postponed are disproportionately risk-averse or health-conscious, those who conceived during the outbreak would be on average less health-conscious than the population of potential mothers, which would tend to overstate the adverse outcomes attributed to reduced prenatal care. Because this selection operates through predetermined maternal characteristics, it is directly testable; we compare observables across exposure levels in Section 3.4 and report formal balance tests, together with a full discussion of instrument validity, in Section 5.2.

[11] One might object that the 3% are unrepresentative of conceptions occurring during the outbreak, since women who conceived during SARS and carried to term delivered in 2004, outside our window. This is correct but immaterial: our estimand is defined over births rather than conceptions, and those deliveries are excluded by the same window that excludes all 2004 births, which leaves the 2003 birth cohort itself unselected.

3.4 Summary Statistics

Our final sample includes approximately 335,000 births in 2002 and 2003, with roughly 40% of observations exposed to the SARS outbreak during pregnancy. Table 1 presents descriptive statistics for the full sample and separately for the unexposed group (deliveries before March 14, 2003) and the exposed group (deliveries on or after March 14, 2003). The two groups exhibit similar daily birth volumes—475.98 and 435.74 births per day, respectively—a difference of roughly 8.5% that reflects differences in calendar coverage across the two periods rather than meaningful compositional differences between samples.

Panel A documents a substantial differential in prenatal care utilization between groups. All three utilization measures—the Kessner adequacy index, the indicator for completing all ten recommended visits, and the total visit count—tell a consistent story: exposed women received significantly less prenatal care than their unexposed counterparts. Adequate care rates fell from 73% in the unexposed group to 49% in the exposed group, consistent with the sharp post-March 2003 decline visible in Figure 2a. The exposed group has a mean SARS exposure fraction of 0.51, indicating that on average these women spent roughly half of their pregnancy during the outbreak period, while the unexposed group has an exposure fraction of 0 by construction.

Despite this large differential in care utilization, raw birth outcomes—including preventable delivery complications, SGA, LBW, and neonatal and infant mortality—appear broadly similar across groups. Rather than being reassuring, this pattern is informative: the absence of detectable differences in raw outcomes despite a large care differential is consistent with opposing selection forces that obscure the true relationship between prenatal care and health outcomes, underscoring the importance of moving beyond raw comparisons to identify the causal effect of prenatal care.

Panel B shows that maternal and birth characteristics are broadly balanced across groups. Mean maternal age is 28.6 years, the male birth share is stable at approximately 52%, and rates of non-preventable complications, including breech presentation, cephalopelvic disproportion, antepartum hemorrhage, placenta previa, cord complications, dystocia, abnormality of organs and soft tissues of pelvis, and previous cesarean delivery that complicates childbirth, are similar across groups.[12] The share of first-time mothers is marginally higher in the exposed group (50.1% vs. 48.6%), a difference of 1.5 percentage points.

The one notable difference between groups is educational attainment: the exposed group exhibits slightly higher educational attainment, 12.95 versus 12.53 years of schooling, and a 3-percentage-point higher share of college graduates. This pattern is unlikely to reflect selective fertility timing in response to SARS (discussed in Section 3.3), since if highly educated women were actively avoiding pregnancy during the outbreak, we would expect the opposite sign. Rather, it reflects the smooth secular increase in educational attainment visible in Figure A3.1a, which shows a continuous upward trend in years of schooling over the full sample period with no de-

[12] We use ICD-9 codes to identify breech presentation, cephalopelvic disproportion, placenta previa, umbilical cord problem, dystocia, abnormality of organs and soft tissues of pelvis, and previous cesarean delivery that complicates childbirth. See Appendix Table A6.1 Panel B for references.

tectable discontinuity at the onset of the shock. We formally assess these patterns, including whether the exposure fraction is uncorrelated with firstborn status, infant sex, maternal age, and nonpreventable complication rates, in the balance tests reported in Table 3 (Section 5.2).

4 Methodology

4.1 OLS Estimation and Biases

To study the impact of decreased prenatal care utilization on maternal and birth outcomes, we begin with the following baseline specification:

$$Y_i = \alpha_0 + \alpha^{OLS} Kessner_i + \boldsymbol{X_i}\pi + \kappa_{age(i)} + \lambda_{m(i)} + \mu_{p(i)} + \nu_{h(i)} + \tau t_i + \epsilon_i, \tag{2}$$

where i indexes births, and the subscripts $age(i)$, $m(i)$, $p(i)$, and $h(i)$ correspond to maternal age, month of birth (from January to December), postal code of residency, and hospital of delivery, respectively. Y_i represents adverse maternal and birth outcomes, including preventable delivery complications, small for gestational age (SGA), low birth weight (LBW), neonatal mortality, and infant mortality. The main explanatory variable, $Kessner_i$, is an indicator for adequate prenatal care following the Kessner Index. $\boldsymbol{X_i}$ is a vector of birth-specific characteristics, including infant sex, birth parity (orders 1, 2, 3, and 4+), maternal education (no formal schooling, elementary, middle, high school, associate, college, and graduate degrees), and a set of nonpreventable birth complications identified via ICD-9 codes. The remaining terms capture maternal age fixed effects ($\kappa_{age(i)}$), birth month fixed effects ($\lambda_{m(i)}$), postal code fixed effects ($\mu_{p(i)}$), hospital fixed effects ($\nu_{h(i)}$), a linear calendar-month time trend (τt_i), and an idiosyncratic error (ϵ_i). Standard errors are clustered at the birth date level.

A critical concern in estimating Equation (2) is omitted variable bias arising from unobserved maternal heterogeneity, such as baseline health status, risk preferences, and health-seeking behavior, that jointly determine prenatal care utilization and birth outcomes. Letting M_i denote this heterogeneity and δ its effect on Y_i, omitting M_i from Equation (2) yields:

$$\operatorname{plim} \hat{\alpha}^{OLS} = \alpha^{TRUE} + \delta \cdot \frac{Cov(Kessner_i, M_i | \boldsymbol{X_i})}{Var(Kessner_i | \boldsymbol{X_i})}. \tag{3}$$

Since Y_i measures adverse outcomes, α^{TRUE} is expected to be negative, reflecting the protective effect of adequate prenatal care. The sign of the bias is a priori ambiguous. On one hand, M_i may capture poor baseline health status. Women with greater health risks are more likely to experience adverse birth outcomes regardless of prenatal care, implying $\delta > 0$. At the same time, these women are more likely to be advised by providers, or motivated by their own condition, to seek more frequent prenatal visits, implying $Cov(Kessner_i, M_i | \boldsymbol{X_i}) > 0$. Together, this produces a positive bias that pushes α^{OLS} toward zero or even into positive territory, understating or reversing the true protective effect of prenatal care. On the other hand, M_i may reflect greater health-

consciousness. Such women are inherently more likely to deliver healthy infants, implying $\delta < 0$, and are simultaneously more likely to adhere to recommended prenatal care guidelines, again implying $Cov(Kessner_i, M_i|\boldsymbol{X_i}) > 0$. This produces a negative bias that makes α^{OLS} more negative than α^{TRUE}, overstating the benefits of prenatal care. Because these two forms of selection push α^{OLS} in opposite directions, and because their relative strength likely varies by outcome, OLS estimates alone cannot reveal either the magnitude or the direction of prenatal care's true effect. We turn next to an instrumental variable strategy that addresses this endogeneity directly.

4.2 2SLS estimation

We implement the instrumental variable approach via two-stage least squares (2SLS) estimation. The first stage takes the form:

$$Kessner_i = b_0 + b_1 SARS_i + \boldsymbol{X_i}\theta + \kappa_{age(i)} + \lambda_{m(i)} + \mu_{p(i)} + \nu_{h(i)} + \tau t_i + \epsilon_i, \tag{4}$$

where $SARS_i$ is the instrument, measured as the fraction of pregnancy exposure to the SARS outbreak (defined in Equation (1)), and all other notation follows Equation (2). The coefficient b_1 captures the effect of SARS exposure on the adequacy of prenatal care. Given the widespread fear of visiting healthcare facilities during the outbreak, we expect $b_1 < 0$—that is, greater SARS exposure leads to fewer prenatal visits and lower care adequacy. As Figure 2a shows, visit counts began declining even before the official start of the outbreak, suggesting that our exposure measure may underestimate the true first-stage effect, making our estimates a lower bound.

The reduced form directly regresses birth outcomes on the instrument:

$$Y_i = c_0 + c_1 SARS_i + \boldsymbol{X_i}\theta + \kappa_{age(i)} + \lambda_{m(i)} + \mu_{p(i)} + \nu_{h(i)} + \tau t_i + \epsilon_i, \tag{5}$$

where c_1 captures the intent-to-treat effect of SARS exposure on birth outcomes. The second stage then replaces $Kessner_i$ with its first-stage predicted value $\widehat{Kessner}_i$, from Equation (4):

$$Y_i = \beta_0 + \beta_1 \widehat{Kessner}_i + \boldsymbol{X_i}\omega + \kappa_{age(i)} + \lambda_{m(i)} + \mu_{p(i)} + \nu_{h(i)} + \tau t_i + \epsilon_i, \tag{6}$$

where β_1 is the coefficient of interest, capturing the causal effect of adequate prenatal care on birth outcomes. The 2SLS estimate β_1 can be interpreted as the ratio of the reduced form to the first stage, c_1/b_1. Standard errors are clustered at the birth date level throughout. We have also estimated models with region-specific time trends, which yield results very similar to those reported below.

For $SARS_i$ to serve as a valid instrument, several conditions must hold. First, **relevance** requires that SARS exposure is correlated with prenatal care adequacy. We expect this to hold given the widespread fear of visiting healthcare facilities during the outbreak, and we verify it formally via the first-stage F-statistic discussed in Section 5.1. Second, **exogeneity** requires that SARS exposure is independent of unobserved determinants of birth outcomes. This is plausible

given that the outbreak was largely unanticipated—over 97% of our sample conceived before the first confirmed case in Taiwan on March 14, 2003—making systematic sorting into exposure unlikely. While two potential selection issues remain regarding the 3% who conceived during the outbreak and women absent from our sample, we discuss in Section 3.3 that both sources of selection bias work in the same direction, if anything, rendering our estimates conservative. Third, the **exclusion restriction** requires that SARS exposure affects birth outcomes only through its effect on prenatal care utilization. Like all IV applications, this assumption is fundamentally untestable. Nevertheless, it is supported by the very low prevalence of SARS infection in our sample (0.07%), which rules out direct health effects of infection as a confounding channel. In Section 6, we conduct a series of empirical tests to rule out alternative channels, including maternal stress and other behavioral responses to the outbreak.

Fourth, **monotonicity** requires that SARS exposure weakly reduces prenatal care adequacy for all women, with no defiers who increase their care in response to the outbreak. This is plausible since the primary mechanism—fear of visiting healthcare facilities—would uniformly discourage prenatal visits, and there is no obvious reason why any subgroup of women would increase their prenatal care utilization as a direct result of SARS exposure. Under this assumption, β_1 should be interpreted as a local average treatment effect (LATE) for compliers—women who reduced their prenatal care utilization as a direct result of SARS exposure. Finally, **no interference (SUTVA)** requires that one woman's exposure does not affect another's potential outcomes. Since $SARS_i$ is defined as the fraction of the 292-day outbreak period overlapping with a woman's own pregnancy, her exposure value is mechanically determined by her own conception and delivery dates relative to the fixed outbreak window, independent of any other individual woman's pregnancy timing or exposure. This construction rules out interference operating through the instrument itself. It does not, on its own, rule out socially-transmitted fear—for instance, a woman skipping visits because women in her social network did (Bennett et al., 2015). Such peer effects, however, would operate through the same channel our design already accommodates: they would affect a woman's own care utilization ($Kessner_i$) rather than her outcomes directly, so they do not threaten the validity of our second-stage estimate. A direct effect of social contagion on outcomes, bypassing care utilization entirely, would instead constitute an exclusion-restriction violation of the kind addressed above.

5 Results

5.1 First Stage and the Treated Margin: Does SARS Exposure Shift Prenatal Care Utilization?

First-Stage Estimates.—Column (1) of Table 2 presents the first-stage estimates, showing a strong relationship between the SARS exposure fraction and prenatal care utilization. In the full sample, moving from zero to complete gestational overlap with the SARS period reduces the probability of adequate prenatal care, as measured by the Kessner index, by 28 percentage points, or 38.5 percent relative to the unexposed mean of 73 percent. Panel B shows nearly identical estimates when the

sample is restricted to full-term births, indicating that the first-stage relationship is not driven by SARS-induced preterm delivery mechanically shortening the time available for prenatal visits.

In Table A4.1, columns (3)-(6), we report robust results under alternative prenatal care measures, namely, the total number of visits and an indicator for receiving at least ten visits. Estimates are virtually unchanged whether or not maternal risk factors, indicators of nonpreventable complications, are controlled for. First-stage F-statistics across all measures range from 1,805 to 2,615, well above conventional thresholds for instrument strength. These first-stage estimates corroborate the visual evidence in Figure 2b, where declining prenatal care adequacy closely tracks the rise in the exposure fraction.

Characterizing the Treated Margin.—To examine the composition of the treated margin and assess whether the estimates are driven by a narrow subpopulation, we investigate treatment effect heterogeneity within a Generalized Random Forest (GRF) framework, following Wager and Athey (2018) and Athey et al. (2019). We estimate Conditional Average Treatment Effects (CATEs) for prenatal care adequacy and partition observations into quintiles according to their predicted responsiveness to the shock. As shown in Figure A5.1, the Augmented Inverse Probability Weighting (AIPW) estimates of realized treatment effects are non-monotonic across these quintiles and statistically indistinguishable from one another, providing little evidence that the GRF's predicted ranking corresponds to systematic differences in realized responsiveness to SARS. Figure A5.2 sheds light on why: among the covariates available to the algorithm, only maternal education varies systematically across quintiles, rising from 11.5 years in the lowest-predicted-responsiveness quintile to 14.8 years in the highest, while maternal age, parity, infant sex, and nonpreventable complications show no clear gradient. This suggests that education is the primary observable characteristic underlying the GRF's ranking, rather than identifying any other observable source of heterogeneity. However, because education was also rising smoothly over our sample period for reasons unrelated to SARS (Appendix A3), and because delivery timing mechanically determines SARS exposure intensity, this education-based ranking likely reflects the secular trend in schooling rather than genuine behavioral heterogeneity in response to the epidemic, consistent with the absence of any corresponding pattern in realized treatment effects.

Taken together, these results provide little evidence that the SARS-induced reduction in prenatal care was concentrated among a narrowly defined subgroup based on the observable characteristics available in our data. Although we cannot establish that compliance was homogeneous or that our LATE equals the Average Treatment Effect, the absence of pronounced heterogeneity in realized first-stage responses distinguishes our setting from instruments whose identifying variation is explicitly concentrated among particular groups, such as eligibility-based expansions in prenatal care access. Unobserved heterogeneity in responsiveness may nevertheless remain.

Timing and Breadth of the Disruption.—Understanding the gestational timing of the disruption is essential for identifying the specific channel through which the epidemic affected prenatal care utilization and, ultimately, health outcomes. Covering 2002–2006, Figure A2.2 plots the probability of achieving trimester-specific care adequacy as a function of birth timing, where trimester-specific

adequacy is defined as the completion of all NHI-subsidized check-ups within a given gestational window: two visits each in the first and second trimesters, and six in the third. The figure shows disruptions throughout pregnancy, with substantially larger declines in second- and third-trimester adequacy, suggesting that SARS affected not only care initiation but also continued engagement with prenatal care later in pregnancy.

To formalize this pattern, we regress trimester-specific adequacy indicators separately on the exposure fraction. Columns (2) through (4) of Panel A in Table 2 report statistically significant declines across all three trimesters, though the magnitude is substantially larger in the later stages of pregnancy. Using the unexposed group—women with zero SARS exposure—as the natural pre-epidemic counterfactual, the likelihood of completing all subsidized visits in the first trimester fell by 4.7 percentage points, representing approximately 6% of the pre-SARS baseline of 75%. The decline is considerably more pronounced in the second trimester, where adequacy fell by 19.8 percentage points—equivalent to 22% of the pre-SARS baseline of 91%—making it the trimester most severely disrupted by the epidemic. The third trimester exhibits a reduction of 3.8 percentage points, or 14% of the pre-SARS baseline of 27%, though the naturally lower baseline adequacy rate in this period warrants careful interpretation, as biological variation in gestational length truncates the available window for late-stage visits even in the absence of any behavioral disruption. Panel B replicates the estimation restricting to full-term births only to rule out the possibility that SARS-induced preterm delivery mechanically reduces measured third trimester adequacy by shortening the gestational window available for late-stage visits. The virtually identical estimates across panels indicate that the decline in third trimester adequacy is not mechanically driven by pregnancies ending earlier due to the shock.

The particularly large second-trimester disruption is clinically relevant because the NHI schedule includes an anomaly scan at approximately week 20, a detailed anatomical examination used to assess fetal development and identify structural abnormalities. If women responded to SARS primarily by forgoing lower-value routine visits while preserving more salient screening services, use of the anomaly scan might have been relatively insulated from the decline in overall prenatal care.

Column (5) of Table 2 examines this possibility. Among unexposed pregnancies, 95 percent received the anomaly scan. Complete gestational exposure to the SARS period reduces scan utilization by 16.7 percentage points, or approximately 18 percent relative to this baseline. Thus, the SARS-induced disruption extended beyond routine checkups to a widely used, clinically important scheduled screening service. This finding suggests that the utilization response was broad rather than confined to lower-salience prenatal visits.

In sum, SARS disrupted prenatal care throughout pregnancy, with disproportionately large reductions in second- and third-trimester care. The pattern therefore reflects not simply delayed or forgone initiation, but reduced continued contact with the prenatal care system as pregnancy progressed. This timing is clinically relevant because prenatal care is structured to generate information at different stages of gestation: some services, such as the second-trimester anomaly scan, are scheduled at particular points, while repeated contact also creates opportunities to identify

conditions that emerge after earlier screening. These patterns motivate the informational mechanism examined in Section 7, where we investigate whether reduced prenatal contact affected the detection of evolving maternal conditions and subsequent delivery management.

5.2 Design Validity: Is the Exposure Fraction Quasi-Randomly Assigned?

Are Maternal and Birth Characteristics Balanced across Exposure Intensity?—Section 3.4 documented that the exposed and unexposed groups are broadly similar across maternal and birth characteristics, with the exception of a small education gradient attributable to secular trends in educational attainment. Here we formally assess instrument validity by regressing five predetermined characteristics—years of schooling, firstborn status, infant sex, maternal age, and nonpreventable complications—on the exposure fraction, using a version of Equation (4) that omits the birth-specific controls X_i, since these characteristics are themselves among the variables included in X_i.[13] Under instrument exogeneity, each coefficient should be statistically indistinguishable from zero, since gestational SARS exposure was determined by delivery date rather than by any deliberate behavioral response to the epidemic.

Table 3 Panel A presents the results. Four of the five coefficients, firstborn, infant sex, maternal age, and nonpreventable complications, are statistically insignificant. Years of schooling is the one statistically significant exception: a unit increase in the exposure fraction is associated with 0.34 additional years of education. This estimate, while statistically significant given our large sample size, is economically small and is consistent with the secular upward trend in educational attainment documented in Sections 3.3 rather than differential selection into SARS exposure. Section A3 further shows that our estimates are insensitive to randomly dropping varying shares of college-educated mothers from the exposed group.

Taken together, the limited association between SARS exposure and these predetermined characteristics is reassuring for the identifying assumption that gestational exposure was not systematically related to observable maternal and pregnancy characteristics. As usual, balance on observables cannot rule out correlation with unobserved determinants of prenatal care or health outcomes; Section 6 therefore provides additional design-based tests of the exclusion restriction.

Is Pre-Pregnancy Health Status Balanced across Exposure Intensity?—A complete validity assessment also requires examining whether women differed systematically in their health status prior to conception, a dimension of selection that cannot be addressed by adding controls, since pre-pregnancy health records are unavailable for our full analysis sample. We therefore use a matched subsample with available pre-pregnancy health records to assess whether such selection, which could otherwise bias our estimates in either direction, is present.

To characterize pre-pregnancy health status, we linked our birth cohort to inpatient records from the two years preceding each delivery. For outpatient records, we matched against three random one-million-beneficiary subsamples of NHI claims, yielding coverage of approximately 14%

[13] We pool all nonpreventable complications into one binary risk indicator because each category is individually rare.

of the cohort.[14] All outcomes are measured over this two-year pre-pregnancy window to avoid conflating baseline health status with pregnancy-induced changes in healthcare utilization.

We consider two classes of outcomes: utilization measures, including any hospitalization, total inpatient days, inpatient expenditures, number of outpatient visits, and outpatient expenditures; and clinical indicators, including antidepressant prescriptions (ATC code N06A),[15] and diagnosed chronic conditions, including hypertension, diabetes, heart disease, and asthma.[16] Within the matched subsample, approximately 1% of women received at least one antidepressant prescription during the observation period, and the average woman had seven outpatient visits annually. The prevalence of chronic conditions was relatively low, reflecting the generally young and healthy population of women of childbearing age.

Table 3 Panel B presents the results. Across both aggregate measures (Panel B(1)) and clinical events (Panel B(2)), the estimated coefficients on the exposure fraction are statistically insignificant and negligible in magnitude, providing no evidence of systematic differences in measured pre-pregnancy health across exposure intensity. Together with the balance results above, these findings provide additional support for the identifying assumptions underlying the IV analysis.

5.3 Selection Bias and Causal Effects: Does Adequate Prenatal Care Improve Maternal and Infant Health?

OLS Estimates and the Direction of Selection Bias.—Panel A of Table 4 reveals a striking contrast across outcome types: adequate prenatal care is associated with higher rates of preventable maternal complications but lower rates of adverse infant outcomes. The framework in Equation (3) illustrates how such patterns can arise from different forms of selection. For maternal complications, the positive OLS coefficients are consistent with provider-driven selection: clinicians may direct women with greater observed or emerging maternal risk toward more intensive monitoring, generating a positive association between prenatal care and underlying complication risk even when care itself is protective. For infant outcomes, by contrast, OLS estimates imply substantial benefits of adequate care, consistent with patient-driven selection if women who seek more prenatal care also engage in other health-promoting behaviors. The OLS estimates imply a 5 percent reduction in SGA and 40–56 percent reductions in neonatal and infant mortality relative to their respective means, effects that disappear under IV below.

Reduced-Form and IV Estimates.—The reduced-form estimates (Panel B) sharply reverse the OLS pattern for maternal outcomes: a 100 percentage point increase in exposure raises the proba-

[14] Each subsample represents roughly 5% of Taiwan's insured population. After pooling the three subsamples and removing duplicates, the matched outpatient sample covers approximately 14% of the primary birth cohort, for whom we observe complete longitudinal healthcare utilization records.

[15] Both outpatient visits and antidepressant prescriptions are likely to be affected by pregnancy itself rather than by pre-existing health: outpatient rise mechanically due to the ten recommended prenatal care visits, and antidepressant use declines substantially during pregnancy (Boone et al., 2025) Measuring both outcomes in the two-year window before conception therefore avoids conflating these gestational-period changes with a woman's underlying health status

[16] Indicators for pre-existing clinical conditions are defined based on ICD diagnoses, including hypertension (ICD-9: 401–405), diabetes (250), heart disease (410–414), and asthma (493).

bility of preventable complications by approximately 0.97–1.19 percentage points, consistent with an adverse effect of reduced prenatal care. For infant outcomes, by contrast, exposure has no statistically detectable effect on any of the four measures.

The 2SLS estimates (Panel C) show that adequate prenatal care reduces the probability of preventable maternal complications by approximately 3.4 percentage points, or 26 percent relative to the sample mean, consistent with the reduced-form results; no statistically significant effects are detected for any infant outcomes. These estimates scale the reduced-form effects by the first-stage relationship between SARS exposure and prenatal care adequacy documented in Section 5.1. The contrast between OLS and IV estimates indicates that observational estimates are biased in opposite directions relative to the IV estimates: they understate the protective effect of prenatal care on maternal complications while overstating its benefits for infant outcomes. We interpret this divergence as consistent with different forms of selection across outcome types.

Robustness.—The main findings are robust across several alternative specifications reported in the appendix. Table A4.2 replaces the Kessner adequacy index with two alternative prenatal care measures—total visit count and an indicator for completing at least ten visits—and finds consistent patterns across all three estimators: OLS understates the benefits for maternal outcomes and overstates them for infant outcomes, while IV estimates concentrate the causal effect on preventable complications. Table A4.3 replaces birth-month fixed effects with conception-month fixed effects to account for seasonal selection at the time of conception. While birth month is known to correlate systematically with maternal characteristics (Buckles and Hungerman, 2013), other studies suggest that seasonality in birth outcomes is more accurately captured by controlling for the month of conception, which better reflects the initial health environment and maternal sorting at the time of pregnancy (Dehejia and Lleras-Muney, 2004; Currie and Schwandt, 2013); point estimates are nearly identical to the baseline throughout, suggesting our main results are not sensitive to how seasonality is controlled. Table A4.4 reports IV probit marginal effects to account for the binary nature of the dependent variables; the estimated marginal effect for preventable complications closely matches the linear probability model estimate, and infant outcome results remain statistically insignificant.

Finally, as documented in Section 3.4, women exposed to SARS during pregnancy have on average slightly higher educational attainment than their unexposed counterparts, reflecting the secular upward trend in schooling rather than endogenous fertility timing. To assess whether this education gradient distorts our findings, we randomly drop 1%, 5%, and 10% of college-educated mothers from the exposed group and re-estimate all specifications (Table A3.1). Point estimates are nearly identical across all three subsamples in both reduced-form and IV specifications, and the pattern of opposing bias across outcome types is fully preserved, indicating that our conclusions are not sensitive to the observed education gradient.

6 Evidence on the Exclusion Restriction

The validity of our instrumental variables strategy rests on the exclusion restriction: that SARS exposure affects maternal and infant outcomes only through its impact on prenatal care utilization, and not through other direct channels. Three such channels are plausible. First, the psychological stress of living through a sudden and lethal epidemic could have harmed maternal and infant health through well-established biological pathways, independently of any reduction in prenatal visits.[17] Second, SARS may have reduced other healthcare utilization beyond prenatal care—outpatient visits, specialist consultations, and elective procedures—in ways that adversely affected maternal or infant health independently of prenatal care[18] Third, pregnant women may have responded to SARS by adopting precautionary and compensatory behaviors—heightened hygiene practices, reduced social contacts, better nutrition, and avoidance of teratogenic substances—that directly improved outcomes independent of prenatal care utilization.[19] The first and second channels would cause our IV estimates to overstate the protective effect of prenatal care, while the third would cause them to understate it. Because these channels operate in opposing directions, the net bias from any exclusion restriction violation is theoretically ambiguous; Appendix A7 formalizes this bias and shows how its direction depends on which channel dominates. We therefore conduct a series of design-based tests that examine distinct empirical predictions of these alternative channels.

Does SARS Exposure Affect Outcomes Prenatal Care Cannot Influence?—Certain outcomes should be unaffected by prenatal care but would respond to direct stress or behavioral channels, making them useful negative controls (Altonji et al., 2005). We examine two such outcomes: infant sex and non-preventable maternal complications. Infant sex is biologically determined at conception and cannot be influenced by prenatal care, but could in principle be affected by stress-induced selective miscarriage (Wu, 2021) or by behavioral changes during pregnancy that differentially affect fetal survival by sex (Azizi et al., 2024; Al-Qaraghouli and Fang, 2017). Non-preventable complications are largely anatomical factors beyond the scope of prenatal care prevention, but could theoretically respond to stress-induced physiological changes such as hormonal fluctuations, altered blood flow, and elevated uterine contractility (Zahran et al., 2013; Gennaro and Hennessy, 2003), or to lifestyle modifications during the outbreak such as reduced physical activity and altered sleep patterns (Yang et al., 2021).

Both the reduced-form effects of SARS exposure and the corresponding 2SLS estimates are statistically indistinguishable from zero for these outcomes (Table A7.1). These null results provide no evidence of direct SARS effects on outcomes that prenatal care itself should not influence, as would be expected under some stress or behavioral channels.

[17] Recent evidence on stress-induced birth complications comes from studies of the COVID-19 pandemic (Liu et al., 2022, 2020). For broader evidence on the adverse effects of maternal stress on birth outcomes, see Aizer et al. (2016) and Persson and Rossin-Slater (2018).

[18] Although direct health effects of SARS infection are unlikely given its very low prevalence—0.07% of the entire population—outpatient visits nonetheless fell by over 30% between March and June of 2003 before returning to pre-SARS levels by September (Bennett et al., 2015), suggesting that indirect disruption to healthcare utilization remains a plausible confounding channel.

[19] See Chang et al. (2004), Chen et al. (2006), Fong and Chang (2011), Wang (2014), and Hsu et al. (2017).

Do First-Time Mothers Show Larger Effects?—If maternal stress were an important direct channel, one might expect larger effects among first-time mothers, who tend to experience greater pregnancy-related anxiety and stress due to uncertainty surrounding pregnancy and childbirth (Gillespie et al., 2018). Table 5 (Panel A) presents IV estimates separately for primiparous and multiparous mothers. Consistent with our main findings, only the estimates for preventable maternal complications are statistically significant for both subgroups. Importantly, the point estimates are very similar across parity groups: $\hat{\beta}_{primipara} = -0.037$ ($SE = 0.012$) versus $\hat{\beta}_{multipara} = -0.033$ ($SE = 0.017$), and a Chow test fails to reject equality of coefficients ($p = 0.83$).[20] The similarity of the estimates across parity groups therefore provides little support for differential maternal stress by parity as the primary explanation for our findings.

Is the Effect Concentrated Among Mothers with Greater Healthcare Needs?—A related concern is that the epidemic reduced general healthcare utilization broadly, with effects on birth outcomes driven by disrupted access to healthcare more generally rather than by reduced prenatal care specifically. The very low prevalence of SARS infection—0.07% of the entire population—makes direct SARS infection unlikely to be an important confounding channel. The more substantive version of this concern operates through indirect disruption: outpatient visits fell by over 30% during the epidemic (Bennett et al., 2015), and mothers with greater underlying healthcare needs, who depend more heavily on continuous medical oversight to manage their health, would be more vulnerable to any general disruption to care access.[21] If this channel were an important driver of our results, we would therefore expect larger effects among mothers with greater baseline healthcare needs. We proxy for healthcare need using prior hospitalization records, on the logic that mothers who were hospitalized in the period before pregnancy carry greater baseline health burdens.

Consistent with our main findings, estimates (Panel B of Table 5) for all outcomes except preventable maternal complications are statistically insignificant for both subgroups. For preventable complications, the point estimates are similar in magnitude across groups, $\hat{\beta}_{hospitalized} = -0.028$ ($SE = 0.049$) and $\hat{\beta}_{non\text{-}hospitalized} = -0.035$ ($SE = 0.011$), and a Chow test fails to reject coefficient equality ($p = 0.89$). The estimate for previously hospitalized mothers is less precise, consistent with the substantially smaller size of this subgroup. We examine heterogeneity more directly by augmenting our main specification with an indicator for prior hospitalization and its interaction with the treatment variable; the interaction term is statistically insignificant across all outcomes (Appendix Table A7.2). We find no evidence of systematically larger effects among mothers with greater pre-pregnancy healthcare needs, as might be expected if broad disruption to non-prenatal healthcare were the primary driver of our results.

Do Direct Channels Operate Where the Instrument Has No Bite?—The East region of Taiwan experienced relatively little SARS-related disruption to prenatal care and showed almost no change

[20] We assess the equality of coefficients across parity groups following the Chow test described in Clogg et al. (1995) and Paternoster et al. (1998).

[21] Outpatient visits returned to pre-SARS levels by September 2003—three months after the last probable case was reported on June 16 (Bennett et al., 2015)—whereas prenatal care utilization exhibited a significantly more protracted recovery trajectory (see Figure A1.2).

in prenatal visit patterns during the outbreak (Figure A1.3). This regional variation provides a useful falsification exercise. We first exclude East-region births from the main sample; estimates remain nearly identical to the baseline, with the coefficient on preventable complications remaining stable at -0.034 ($SE = 0.01$) (Table 5 Panel C(1)). We next examine the East region separately. If gestational SARS exposure affects outcomes primarily through its effect on prenatal care, a region in which exposure has little effect on prenatal care should also exhibit little reduced-form effect on health outcomes. Consistent with this prediction, the East-region first stage is small and statistically insignificant ($\hat{b}_1 = -0.048$, $SE = 0.031$, $p = 0.12$; Appendix Table A4.1, Column (8)), and reduced-form estimates are statistically insignificant across the primary outcomes (Panel C(2)). This "double-zero" pattern is consistent with the exclusion restriction and provides no evidence of sizeable direct effects of SARS exposure in a region where its effect on prenatal care was minimal.

Taken together, these tests examine distinct implications of the leading alternative channels through which SARS exposure could directly affect maternal or infant health. We find no effects on negative-control outcomes that could respond to stress or behavioral changes, no systematically larger effects among first-time mothers or women with greater pre-pregnancy healthcare needs, and no detectable reduced-form effects in the East region, where SARS exposure had little effect on prenatal care. Although no individual test can establish the exclusion restriction, the consistency of these findings across distinct tests provides reassurance that the main IV estimates are unlikely to be driven primarily by direct effects of the epidemic.

7 Mechanism: Through What Channels Does Adequate Prenatal Care Operate?

The main results establish that adequate prenatal care significantly reduces preventable maternal complications. We now examine potential channels through which this effect operates. We argue that prenatal care functions, in part, as a longitudinal information system: repeated visits generate clinical information that allows providers to detect evolving maternal risks and incorporate that information into subsequent care and delivery decisions. We present evidence in three parts. First, we examine whether adequate care affects the recorded diagnosis of anemia, a condition whose detection may require additional testing later in pregnancy, while having little effect on a condition screened through routine measurement at every visit. Second, we examine whether adequate care affects delivery-mode decisions and assess how much of the headline effect could be accounted for mechanically by the resulting shift in delivery mode. Third, we examine whether preventable complications also decline among women who ultimately deliver vaginally. Each piece of evidence is subject to its own caveats, which we discuss in turn; together, they provide a partial account of how adequate prenatal care may reduce preventable maternal complications.

Detection: Missed Visits and Missed Diagnoses.—We examine the detection channel by contrasting two conditions that are often asymptomatic in their mild forms but differ in what a prenatal visit must entail for routine detection. Anemia is diagnosed through laboratory mea-

surement of hemoglobin or hematocrit.[22] Hypertensive disorders of pregnancy, by contrast, can be detected through routine blood pressure measurement at prenatal visits, with elevated blood pressure prompting further evaluation for proteinuria or other evidence of preeclampsia.[23]

This distinction is particularly useful in the institutional setting we study. As documented in the contemporaneous Maternal Health Booklet reproduced in Appendix Figure A2.1, blood pressure and urine testing for protein and glucose were components of the "routine prenatal test" performed at each scheduled checkup. Thus, every attended visit provided an opportunity to detect elevated blood pressure without requiring either recognizable symptoms or a discretionary diagnostic decision to order a separate test. Hemoglobin and hematocrit testing, by contrast, were included in the laboratory panel at the first prenatal visit but were not part of the routine test at subsequent checkups. Detecting anemia that developed later in pregnancy therefore required additional blood testing that was not automatically performed at each visit. Continued contact with providers created opportunities for such testing, whereas missed visits reduced those opportunities.

A second feature of anemia reinforces this distinction: anemia is not necessarily detectable from the initial prenatal blood test. Maternal plasma volume expansion causes hemoglobin concentrations to decline over pregnancy, reaching their nadir in the second trimester (Ubom et al., 2025). A woman may therefore screen normal early in pregnancy and develop anemia subsequently. Consistent with this progression, first-trimester hemoglobin predicts third-trimester anemia status with 83 percent sensitivity (Noshiro et al., 2022), indicating that an early normal screen does not identify all women who will be anemic later in pregnancy. Detecting anemia that develops subsequently requires later testing and may therefore depend on continued contact with the healthcare system.

These differences yield a direct empirical prediction. If missed prenatal visits reduce opportunities for additional diagnostic testing, recorded anemia should decline because detecting cases that develop after the initial screen requires subsequent testing. This mechanism should be less important for hypertensive disorders, for which blood pressure measurement is incorporated into every routine prenatal visit and does not require a separate laboratory order.

Table 6 supports this prediction. Anemia (column 1) declines by 0.81 percentage points ($\hat{\beta} = -0.0081$, SE $= 0.0016$), a 31 percent reduction relative to its 2.6 percent baseline mean. Hypertensive disorders of pregnancy (column 2), in contrast, show no statistically detectable change ($\hat{\beta} = -0.0009$, SE $= 0.0012$; mean $= 0.013$). This asymmetry is consistent with the institutional structure of prenatal monitoring and suggests that the effect of missed visits on diagnosis depends not simply on whether a condition is asymptomatic, but on whether its detection requires additional testing that may be triggered by continued contact with providers. Because our data capture recorded diagnoses rather than underlying disease prevalence, these estimates cannot distinguish missed detection from any direct effect of prenatal care on the incidence of anemia. The differential

[22] Mild anemia in pregnancy is often asymptomatic or produces only nonspecific symptoms, such as fatigue and weakness, that are common in normal pregnancy and therefore provide little diagnostic signal absent laboratory testing.

[23] Gestational hypertension and preeclampsia without severe features may be asymptomatic and are often first identified through routine blood-pressure measurement; suspected preeclampsia may then prompt additional evaluation for proteinuria and other evidence of maternal organ dysfunction (National Institute for Health and Care Excellence, 2019).

monitoring requirements across the two conditions, however, make reduced detection a plausible interpretation of the observed pattern. Moreover, to the extent that prenatal care prevents or treats anemia, any reduction in underlying anemia prevalence would work against the positive effect of care on recorded diagnosis.

Delivery-Mode Decisions: A Mechanical Benchmark.—Prenatal monitoring does more than detect maternal complications; it also generates clinical information that can affect delivery planning. Some indications for cesarean delivery emerge or can only be confirmed later in pregnancy, so gaps in prenatal monitoring may leave these risks unidentified before labor.[24] Delivery-mode decisions are a well-documented margin on which obstetric decision quality affects patient outcomes: Currie and MacLeod (2017) shows that variation in physicians' decision-making translates into differences in maternal and infant complications, independent of procedural skill. Consistent with this mechanism, column (3) of Table 6 shows that SARS exposure reduces the probability of cesarean delivery by 2.0 percentage points (mean = 0.344). The corresponding IV estimate indicates that adequate prenatal care increases the probability of cesarean delivery by 7.2 percentage points.

The magnitude of this delivery-mode response raises the question of how much of the headline effect on preventable complications could arise mechanically from the resulting change in the composition of deliveries. Preventable complications, as defined in Section 3.2, are overwhelmingly concentrated among vaginal deliveries: their prevalence is 19.6 percent among vaginal births, compared with only 0.6 percent among cesarean births (N = 220,275 and 115,395, respectively). As a mechanical benchmark, holding these within-mode complication rates fixed, reallocating 7.2 percentage points of births from vaginal to cesarean delivery would reduce the population complication rate by $0.072 \times (0.196 - 0.006) \approx 0.014$, or approximately 1.4 percentage points. This corresponds to roughly 40 percent of the 3.44 percentage-point headline effect. The calculation is not a causal mediation decomposition; rather, it shows that the observed delivery-mode response is quantitatively large enough to account for roughly 40 percent of the magnitude of the overall reduction in preventable complications under this mechanical benchmark.

We next ask whether the data also show lower preventable complication rates among women who ultimately deliver vaginally. Because delivery mode is itself affected by prenatal care, these within-mode estimates do not have the same causal interpretation as our main IV estimates: conditioning on realized delivery mode may change the composition of women in each subsample. With this caveat in mind, we re-estimate the main specification separately by realized delivery mode. Among vaginal deliveries, adequate prenatal care is associated with a 3.86 percentage-point reduction in preventable complications in the IV specification (Appendix Table A8.1), an estimate similar to the 3.44 percentage-point effect in the full sample. Among cesarean deliveries, the corresponding

[24] Some relevant risk factors are not fixed facts ascertainable at a single early visit but require confirmation later in pregnancy. Fetal malpresentation is a leading example: the majority of breech fetuses identified earlier in pregnancy turn spontaneously to a cephalic position before term, so presentation is not reliably known until later in pregnancy (Salim et al., 2021). In a UK cohort study, introducing routine late third-trimester scanning reduced the share of breech presentations undiagnosed before labor from 22.3 percent to 4.7 percent (Salim et al., 2021). Absent late-pregnancy assessment, a substantial share of breech presentations may therefore remain unrecognized heading into labor.

estimate is statistically indistinguishable from zero (Appendix Table A8.1), and equality of the two within-mode estimates is rejected ($p = 0.02$).

These within-mode estimates should not be interpreted as components that can be added to, or subtracted from, the 40 percent mechanical benchmark above. In particular, the vaginal-delivery estimate does not identify the "remaining" effect after removing the delivery-mode channel. Rather, it provides complementary evidence that the overall pattern is not confined to shifts between vaginal and cesarean delivery. Subject to the selection induced by conditioning on realized delivery mode, the vaginal-delivery result is consistent with an additional within-mode channel through which information accumulated during prenatal care may help providers anticipate and manage risks during labor and delivery.

Taken together, these results are consistent with prenatal care improving maternal outcomes partly through the accumulation and use of clinical information. Continued prenatal contact increases opportunities to detect evolving risks, appears to affect delivery-mode decisions, and may also facilitate risk management within delivery mode.

8 Conclusions

This paper uses Taiwan's 2003 SARS epidemic to estimate the causal effect of prenatal care on maternal and infant health. Exploiting differences across pregnancies in gestational exposure to the sharp and unanticipated SARS-induced disruption to prenatal care, we find that adequate care reduces preventable maternal delivery complications by 3.4 percentage points, or 26 percent, while having no statistically detectable effect on the infant outcomes we examine. A series of design-based tests provides no evidence that direct effects of SARS, rather than disrupted prenatal care, account for these findings.

Beyond these average effects, our results highlight two features of prenatal care that may help reconcile seemingly conflicting findings in the existing literature. First, the direction of selection appears to differ across outcomes. The opposite-signed divergence between OLS and IV estimates for maternal and infant outcomes is consistent with competing patient- and provider-driven sources of selection, whose relative importance may depend on what underlying risks are observable during pregnancy. Second, prenatal care appears to matter not simply through the services provided at individual visits, but through the clinical information accumulated across pregnancy. Our mechanism evidence is consistent with repeated monitoring creating opportunities to detect evolving risks and allowing accumulated clinical information to inform subsequent delivery decisions. Together, these findings shift attention from prenatal care as a simple measure of healthcare quantity toward the processes of selection, information generation, and clinical decision-making that underlie observed relationships between care utilization and health.

These findings have implications beyond the SARS episode. Evaluations of prenatal care—and health investments more generally—should not presume that observational selection operates uniformly across outcomes. When treatment intensity reflects both patient care-seeking and provider

responses to clinical risk, even the direction of selection may vary with the outcome being studied. Our findings also highlight the informational value of continuity in medical care. Interruptions in repeated contact can reduce the clinical information available for subsequent decisions, even when patients continue to receive some care. This distinction may matter during pandemics, natural disasters, capacity constraints, and other disruptions to routine care: preserving clinically important monitoring may be valuable not only for the services delivered during those encounters, but also for the information they generate for later decisions.

Our results should be interpreted with several limitations in mind. As with any instrumental-variables design, the exclusion restriction cannot be tested directly. However, the balance tests, negative-control outcomes, subgroup comparisons, and regional placebo exercises provide no evidence that the leading alternative channels account for our findings. Our IV estimates are local to women whose prenatal care responded to SARS-induced disruption. Although we find little evidence that the first-stage response was concentrated among a narrowly defined subgroup based on observable maternal characteristics, unobserved heterogeneity in responsiveness may remain. Finally, Taiwan during this period had near-universal insurance coverage and high baseline prenatal-care utilization, so the magnitude of the effects may not generalize to settings in which financial barriers are substantial or baseline access is much lower.

More broadly, our findings suggest that asking whether prenatal care "works" may be too coarse a question. Understanding its effects requires considering the outcomes being studied, the selection processes governing care utilization, and the information generated through continued clinical contact. Prenatal care is therefore not simply a sequence of checkups: it is an ongoing process through which patients and providers respond to evolving risks and accumulate information that can shape subsequent clinical decisions. Recognizing this broader role may help explain why estimates of prenatal care effects differ across outcomes and studies.

References

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

Aizer, A., Stroud, L. and Buka, S. (2016), 'Maternal stress and child outcomes: Evidence from siblings', *Journal of Human Resources* **51**(3), 523–555.

Al-Qaraghouli, M. and Fang, Y. M. V. (2017), 'Effect of fetal sex on maternal and obstetric outcomes', *Frontiers in Pediatrics* **5**, 144.

Altonji, J. G., Elder, T. E. and Taber, C. R. (2005), 'Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools', *Journal of Political Economy* **113**(1), 151–184.

Athey, S., Tibshirani, J. and Wager, S. (2019), 'Generalized random forests', *The Annals of Statistics* **47**(2), 1148–1178.

Azizi, B., Najafi, E., Aghamohammadi, V., Nasiri, K., Moghadam, H. R., Bazyar, H., Hosseini, A., Sharifi, E. and Salari, F. (2024), 'Evaluation of the impact of environmental pollutants on the sex ratio: a systematic review', *Reviews on Environmental Health* .

Bennett, D., Chiang, C.-F. and Malani, A. (2015), 'Learning during a crisis:the SARS epidemic in Taiwan', *Journal of Development Economics* **112**, 1–18.

Bjerk, D. J. (2025), Understanding IV versus OLS estimates of treatment effects and the coefficient difference check, IZA Discussion Paper 18274, Institute of Labor Economics (IZA).

Black, D. A., Joo, J., LaLonde, R. J., Smith, J. A. and Taylor, E. J. (2022), 'Simple tests for selection: Learning more from instrumental variables', *Labour Economics* **79**, 102237.

Boone, C., Colina, C. and Pope, D. (2025), 'Antidepressant use before, during, and after pregnancy', *JAMA Network Open* **8**(1), e2457324.

Buckles, K. S. and Hungerman, D. M. (2013), 'Season of birth and later outcomes: Old questions, new answers', *Review of Economics and Statistics* **95**(3), 711–724.

Campbell, O. M. R. and Graham, W. J. (2006), 'Strategies for reducing maternal mortality: Getting on with what works', *The Lancet* **368**(9543), 1284–1299.
URL: *https://doi.org/10.1016/S0140-6736(06)69381-1*

Chang, H.-J., Huang, N., Lee, C.-H., Hsu, Y.-J., Hsieh, C.-J. and Chou, Y.-J. (2004), 'The impact of the sars epidemic on the utilization of medical services: Sars and the fear of sars', *American Journal of Public Health* **94**(4), 562–564.

Chen, T.-J., Lin, M.-H., Chou, L.-F. and Hwang, S.-J. (2006), 'Hospice utilization during the sars outbreak in taiwan', *BMC Health Services Research* **6**(1), 1–9.

Ciacci, R. (2025), 'A matter of size: Comparing IV and OLS estimates', *PLOS ONE* **20**(10), e0334392.

Clogg, C. C., Petkova, E. and Haritou, A. (1995), 'Statistical methods for comparing regression coefficients between models', *American Journal of Sociology* **100**(5), 1261–1293.

Corman, H., Dave, D. and Reichman, N. E. (2019), The effects of prenatal care on birth outcomes: Reconciling a messy literature, *in* 'Oxford Research Encyclopedia of Economics and Finance'.

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

Currie, J. and MacLeod, W. B. (2017), 'Diagnosing expertise: Human capital, decision making, and performance among physicians', *Journal of labor economics* **35**(1), 1–43.

Currie, J. and Schwandt, H. (2013), 'Within-mother analysis of seasonal patterns in health at birth', *Proceedings of the National Academy of Sciences* **110**(30), 12265–12270.

Cygan-Rehm, K. and Karbownik, K. (2022), 'The effects of incentivizing early prenatal care on infant health', *Journal of Health Economics* **83**, 102612.

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

Einav, L., Finkelstein, A., Oostrom, T., Ostriker, A. and Williams, H. (2020), 'Screening and selection: The case of mammograms', *American Economic Review* **110**(12), 3836–3870.

Evans, W. N. and Lien, D. S. (2005), 'The benefits of prenatal care: evidence from the pat bus strike', *Journal of Econometrics* **125**(1-2), 207–239.

Fong, E. and Chang, L.-y. (2011), 'Community under stress: Trust, reciprocity, and community collective efficacy during sars outbreak', *Journal of Community Health* **36**(5), 797–810.

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

Gennaro, S. and Hennessy, M. D. (2003), 'Psychological and physiological stress: impact on preterm birth', *Journal of Obstetric, Gynecologic, & Neonatal Nursing* **32**(5), 668–675.

Gillespie, S. L., Mitchell, A. M., Kowalsky, J. M. and Christian, L. M. (2018), 'Maternal parity and perinatal cortisol adaptation: The role of pregnancy-specific distress and implications for postpartum mood', *Psychoneuroendocrinology* **97**, 86–93.

Hsu, Y.-C., Chen, Y.-L., Wei, H.-N., Yang, Y.-W. and Chen, Y.-H. (2017), 'Risk and outbreak communication: lessons from taiwan's experiences in the post-sars era', *Health Security* **15**(2), 165–169.

Jewell, R. T. and Triunfo, P. (2006), 'The impact of prenatal care on birthweight: the case of uruguay', *Health Economics* **15**(11), 1245–1250.

Kessner, D. M., Singer, J., Kalk, C. E. and Schlesinger, E. R. (1973), *Infant Death: An Analysis by Maternal Risk and Health Care*, Institute of Medicine, National Academy of Sciences, Washington, DC.

Knobler, S., Mahmoud, A., Lemon, S., Mack, A., Sivitz, L., Oberholtzer, K. et al. (2004), The public health response to sars, *in* 'Learning from SARS: Preparing for the Next Disease Outbreak: Workshop Summary', National Academies Press (US).

Landon, M. B., Hauth, J. C., Leveno, K. J., Spong, C. Y., Leindecker, S., Varner, M. W., Moawad, A. H., Caritis, S. N., Harper, M., Wapner, R. J. et al. (2005), 'Maternal and perinatal outcomes associated with a trial of labor after prior cesarean delivery', *Obstetrical & Gynecological Survey* **60**(5), 291–293.

Lautharte, I. (2021), 'Babies and bandidos: Birth outcomes in pacified favelas of Rio de Janeiro', *Journal of Health Economics* **77**, 102457.

Lee, J.-W., McKibbin, W. J. et al. (2004), Estimating the global economic costs of sars, *in* 'Learning from SARS: preparing for the next disease outbreak: workshop summary', National Academies Press Washington, DC, pp. 92–109.

Liu, C. H., Hyun, S., Erdei, C. and Mittal, L. (2022), 'Prenatal distress during the covid-19 pandemic: clinical and research implications', *Archives of Gynecology and Obstetrics* **306**(2), 397–405.

Liu, X., Chen, M., Wang, Y., Sun, L., Zhang, J., Shi, Y., Wang, J., Zhang, H., Sun, G., Baker, P. N. et al. (2020), 'Prenatal anxiety and obstetric decisions among pregnant women in wuhan and chongqing during the covid-19 outbreak: a cross-sectional study', *BJOG: An International Journal of Obstetrics & Gynaecology* **127**(10), 1229–1240.

Maunder, R. (2004), 'The experience of the 2003 sars outbreak as a traumatic stress among frontline healthcare workers in toronto: lessons learned', *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**(1447), 1117–1125.

National Institute for Health and Care Excellence (2019), Hypertension in pregnancy: diagnosis and management, NICE Guideline NG133, NICE. Published: 25 June 2019. Accessed: 22 April 2026.
URL: *https://www.nice.org.uk/guidance/ng133*

Nickell, L. A., Crighton, E. J., Tracy, C. S., Al-Enazy, H., Bolaji, Y., Hanjrah, S., Hussain, A., Makhlouf, S. and Upshur, R. E. (2004), 'Psychosocial effects of sars on hospital staff: survey of a large tertiary care institution', *Cmaj* **170**(5), 793–798.

Noshiro, K., Umazume, T., Hattori, R., Kataoka, S., Yamada, T. and Watari, H. (2022), 'Hemoglobin concentration during early pregnancy as an accurate predictor of anemia during late pregnancy', *Nutrients* **14**(4), 839.

Oberholtzer, K., Sivitz, L., Mack, A., Lemon, S., Mahmoud, A., Knobler, S. et al. (2004), 'Learning from sars: preparing for the next disease outbreak: workshop summary'.

Oster, E. (2020), 'Health recommendations and selection in health behaviors', *American Economic Review: Insights* **2**(2), 143–160.

Paternoster, R., Brame, R., Mazerolle, P. and Piquero, A. (1998), 'Using the correct statistical test for the equality of regression coefficients', *Criminology* **36**(4), 859–866.

Persson, P. and Rossin-Slater, M. (2018), 'Family ruptures, stress, and the mental health of the next generation', *American Economic Review* **108**(4-5), 1214–52.

Salim, I., Staines-Urias, E., Mathewlynn, S., Drukker, L., Vatish, M. and Impey, L. (2021), 'The impact of a routine late third trimester growth scan on the incidence, diagnosis, and management of breech presentation in oxfordshire, UK: A cohort study', *PLoS Medicine* **18**(1), e1003503.

Silverman, M. and Turrentine, M. A. (2023), 'Urinary tract infections in pregnant individuals', *Obstet Gynecol* **142**(2), 435–45.

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

Centers for Disease Control and Prevention (CDC) (2003), 'Severe acute respiratory syndrome–taiwan, 2003', *MMWR. Morbidity and Mortality Weekly Report* **52**(20), 461–466.

Ubom, A. E., Begum, F., Ramasauskaite, D., Nieto-Calvache, A. J., Oguttu, M., Nunes, I., Malel, Z. J., Beyeza-Kashesya, J. and Wright, A. (2025), 'FIGO good practice recommendations on anemia in pregnancy, to reduce the incidence and impact of postpartum hemorrhage (PPH)', *International Journal of Gynecology & Obstetrics* **171**(3), 993–1007.

US Preventive Services Task Force (2018), 'Screening for syphilis infection in pregnant women: Us preventive services task force reaffirmation recommendation statement', *Jama* **320**(9), 911–917.

Wager, S. and Athey, S. (2018), 'Estimation and inference of heterogeneous treatment effects using random forests', *Journal of the American Statistical Association* **113**(523), 1228–1242.

Wang, K.-Y. (2014), 'How change of public transportation usage reveals fear of the sars virus in a city', *PloS One* **9**(3), e89405.

World Health Organization (2003), 'Cumulative number of reported probable cases of severe acute respiratory syndrome (SARS)'.

World Health Organization (2016), *WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience*, World Health Organization, Geneva.
URL: *https://www.who.int/publications/i/9789241549912*

Wu, H. (2021), 'Maternal stress and sex ratio at birth in sweden over two and a half centuries: a retest of the trivers–willard hypothesis', *Human Reproduction* **36**(10), 2782–2792.

Yan, J. (2017), 'The effects of prenatal care utilization on maternal health and health behaviors', *Health Economics* **26**(8), 1001–1018.

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

Yang, Y., Li, W., Yang, W., Wang, L., Liu, J., Leng, J., Li, W., Wang, S., Li, J., Hu, G. et al. (2021), 'Physical activity and sleep duration during pregnancy have interactive effects on caesarean delivery: a population-based cohort study in tianjin, china', *BMC Pregnancy and Childbirth* **21**(1), 406.

Zahran, S., Peek, L., Snodgrass, J. G., Weiler, S. and Hempel, L. (2013), 'Abnormal labor outcomes as a function of maternal exposure to a catastrophic hurricane event during pregnancy', *Natural Hazards* **66**(1), 61–76.

Figures

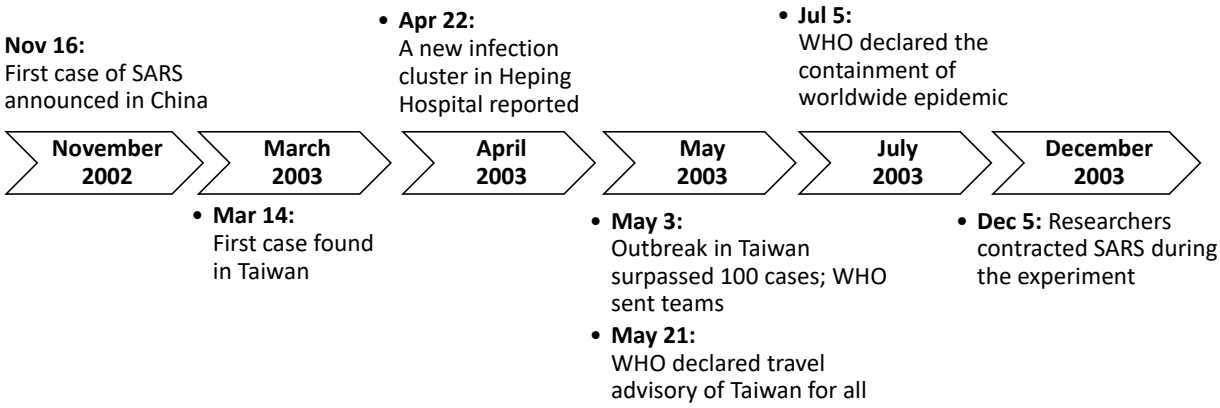
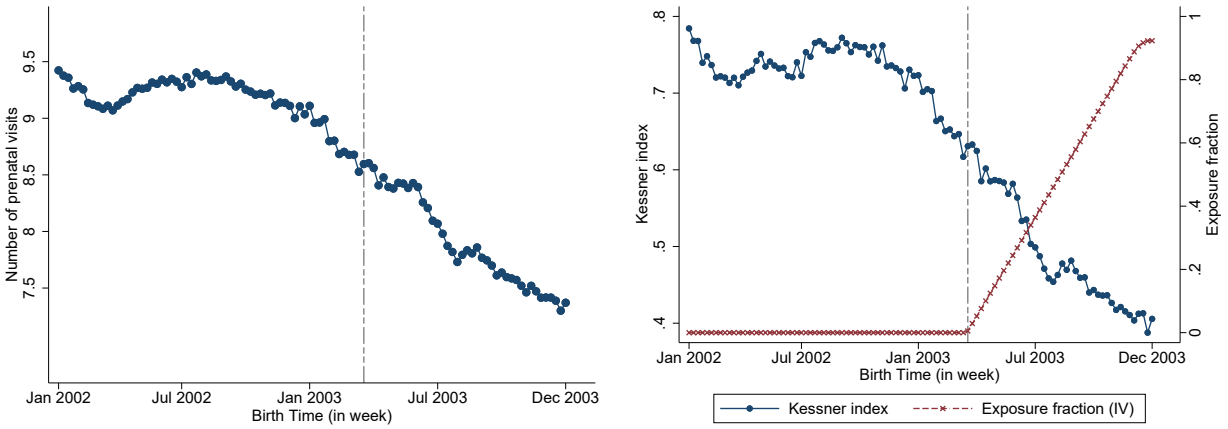


Figure 1: Key SARS Epidemic Milestones in Taiwan, 2002–2003

Notes: The timeline traces seven key events: SARS emergence in China (Nov. 16, 2002); Taiwan's first confirmed case and the start of our treatment period (Mar. 14, 2003); the Heping Hospital nosocomial cluster and declaration of a national public health emergency (Apr. 22, 2003); Taiwan surpassing 100 cumulative cases and WHO dispatch of investigative teams (May 3, 2003); the WHO travel advisory for Taiwan (May 21, 2003); WHO's declaration of worldwide containment and Taiwan's removal from the affected-areas list (Jul. 5, 2003); and a laboratory SARS infection of researchers in Taiwan (Dec. 5, 2003), underscoring the persistence of public anxiety beyond the formal end of the epidemic.

Figure 2: Prenatal Care Utilization and SARS Exposure Fraction by Birth Time in Weeks



(a) The Number of Prenatal Visits

(b) Prenatal Adequacy and SARS Exposure

Notes: Appendix Figure A1.2 plots the number of prenatal visits covering 2002–2006, showing that prenatal care utilization did not recover to pre-epidemic levels until after 2006. In Figure 2b, prenatal adequacy is defined as an indicator defined by the Kessner adequacy index; the IV "exposure" is defined as the fraction of the pregnancy duration that overlaps with the SARS outbreak, normalized by the total duration of the outbreak.

Tables

Table 1: Summary Statistics

	(1) Full sample	(2) No exposure	(3) Any exposure
Panel A. Prenatal care utilization, IV, and outcome variables			
Kessner adequacy index	63.73%	72.95%	48.70%
10 prenatal care visits	49.29%	56.95%	36.81%
Number of prenal care visits	8.68	9.16	7.88
Exposure fraction (IV)	0.19	0	0.51
Preventable complications	13.06%	13.69%	12.03%
Small for gestational age	11.76%	11.70%	11.86%
Low birth weight	4.86%	4.92%	4.75%
Neonatal mortality per 1,000	1.41	1.50	1.27
Infant mortality per 1,000	3.06	3.14	2.92
Panel B. Maternal and birth characteristics			
Age at delivery	28.58	28.51	28.70
15-19	2.46%	2.61%	2.21%
20-24	17.38%	17.76%	16.77%
25-29	37.87%	37.80%	38.00%
30-34	31.47%	31.25%	31.85%
35-40	10.81%	10.58%	11.18%
Maternal education and attainment			
Years of schooling	12.69	12.53	12.95
No school	0.03%	0.03%	0.02%
Elementary school	1.30%	1.60%	0.81%
Junior high school	13.12%	15.06%	9.97%
Senior high school	45.34%	45.54%	45.01%
Associate degree	24.43%	23.26%	26.33%
Bachelor degree	13.45%	12.43%	15.11%
Graduate degree	2.33%	2.07%	2.74%
Infant gender (male)	52.36%	52.26%	52.51%
Firstborn	49.15%	48.55%	50.13%
Parity	1.66	1.67	1.64
Nonpreventable complications			
Breech	8.55%	8.63%	8.41%
Cephalopelvic disproportion	3.27%	3.32%	3.19%
Placenta previa	1.40%	1.40%	1.39%
Cord complication	1.02%	1.03%	1.00%
Dystocia	7.36%	7.28%	7.48%
Abnormality of organs and soft tissues of pelvis	0.02%	0.02%	0.02%
Previous Cesarean Delivery	14.43%	14.66%	14.05%
Births per day	459.83	475.98	435.74
Number of observations	335,674	208,002	127,672

Notes: We use ICD-9 codes to identify breech presentation, cephalopelvic disproportion, antepartum hemorrhage, placenta previa, umbilical cord complication, dystocia, abnormality of organs and soft tissues of pelvis, and previous cesarean delivery that complicates childbirth. See Appendix Table A6.1 Panel B for references.

Table 2: Reduction in Kessner and Trimester-Specific Care Adequacy and Second-Trimester Anomaly Scan

	(1) Kessner Adequacy	(2) Adequacy 1st Trimester	(3) Adequacy 2nd Trimester	(4) Adequacy 3rd Trimester	(5) 2nd-Trimester Anomaly Scan
Panel A. Full Sample					
Exposure fraction	-0.2812***	-0.0467***	-0.1977***	-0.0383***	-0.1673***
	(0.0061)	(0.0074)	(0.0043)	(0.0070)	(0.0040)
Mean of dep var	0.638	0.741	0.855	0.240	0.903
Pre-period dep var mean	0.730	0.746	0.908	0.268	0.954
Observations	332,354	332,354	332,354	332,354	332,354
Panel B. Full-Term Births					
Exposure fraction	-0.2855***	-0.0478***	-0.1986***	-0.0417***	-0.1655***
	(0.0062)	(0.0076)	(0.0044)	(0.0073)	(0.0040)
Mean of dep var	0.655	0.739	0.855	0.254	0.908
Pre-period dep var mean	0.749	0.744	0.907	0.284	0.956
Observations	312,629	312,629	312,629	312,629	312,629
Risk factors	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes
Birth month FE	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: We define trimester adequacy as the completion of all NHI-subsidized check-ups within a given period: two visits in both the first and second trimesters, and six in the third. We note that the baseline probability of adequacy is naturally lower in the third trimester due to biological variation in gestational length. Second-trimester anomaly scan is scheduled at around week 20 according to the NHI Maternal Health Booklet.

Table 3: Balance Tests and Treatment Effects on Pre-Pregnancy Healthcare Utilization

	(1)	(2)	(3)	(4)	(5)
Panel A. Balance test					
	Years of Schooling	Firstborn	Male Infant	Age	Nonpreventable Complications
Exposure fraction	0.3442***	0.0130	-0.0039	-0.0914	-0.0169
	(0.0250)	(0.0070)	(0.0055)	(0.0630)	(0.0151)
Mean of dep var	12.691	0.491	0.524	28.581	0.302
Observations	335,674	335,674	335,674	335,674	335,674
Panel B. Pre-pregnancy healthcare utilization and health status					
(1) Aggregate measures					
	Hospitalization	Length of Stay	Inpatient Exp	Outpatient Visits	Outpatient Exp
Exposure fraction	0.0014	-0.0048	-26.5828	-0.3026	-155.7970
	(0.0025)	(0.0217)	(85.4417)	(0.2261)	(124.0089)
Mean of dep var	0.053	0.257	1249.267	7.218	3238.912
Observations	335,674	335,674	335,674	45,707	45,707
(2) Clinical events					
	Antidepressant Prescription	Hypertension	Diabetes	Heart Disease	Asthma
Exposure fraction	-0.0008	0.0001	-0.0004	-0.0007	-0.0027
	(0.0029)	(0.0015)	(0.0016)	(0.0010)	(0.0031)
Mean of dep var	0.010	0.003	0.003	0.001	0.011
Observations	45,707	45,707	45,707	45,707	45,707
Risk factors	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes
Birth month FE	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Estimated Effects of Prenatal Care on Maternal and Birth Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Preventable Complications		SGA		LBW		Neonatal Mortality		Infant Mortality	
Panel A. OLS estimation										
Kessner adequacy	0.0113***	0.0095***	-0.0066***	-0.0063***	-0.0446***	-0.0438***	-0.8018***	-0.7903***	-1.2716***	-1.2512***
	(0.0011)	(0.0010)	(0.0013)	(0.0013)	(0.0010)	(0.0010)	(0.1604)	(0.1601)	(0.2373)	(0.2377)
Panel B. Reduced-form estimation										
Exposure fraction	0.0119**	0.0097**	0.0061	0.0061	-0.0033	-0.0027	0.5216	0.5474	0.9515	0.9833
	(0.0041)	(0.0030)	(0.0036)	(0.0035)	(0.0023)	(0.0023)	(0.3846)	(0.3855)	(0.6152)	(0.6151)
Panel C. IV estimation										
Kessner adequacy	-0.0423**	-0.0344**	-0.0216	-0.0218	0.0116	0.0094	-1.8528	-1.9413	-3.3798	-3.4874
	(0.0147)	(0.0107)	(0.0127)	(0.0126)	(0.0080)	(0.0081)	(1.3651)	(1.3662)	(2.1888)	(2.1852)
Risk factors	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	0.131	0.131	0.118	0.118	0.049	0.049	1.412	1.412	3.057	3.057
Observations	335,674	335,674	335,674	335,674	335,674	335,674	335,674	335,674	335,674	335,674

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Effects of Prenatal Care Adequacy: Tests of the Exclusion Restriction

	(1) Prev. Complications	(2) SGA	(3) LBW	(4) Neonatal Mortality	(5) Infant Mortality
Panel A. Stratification by firstborn					
(1) Primiparas (First-time mothers)					
Kessner adequacy	-0.0373**	-0.0183	0.0145	-2.4266	-3.3825
	(0.0123)	(0.0157)	(0.0107)	(1.9522)	(3.0577)
Mean of dep var	0.099	0.095	0.044	1.435	3.275
Observations	170,700	170,700	170,700	170,700	170,700
(2) Multiparas (Experienced mothers)					
Kessner adequacy	-0.0327*	-0.0244	0.0064	-1.4529	-3.4650
	(0.0166)	(0.0181)	(0.0120)	(1.9967)	(3.0315)
Mean of dep var	0.164	0.141	0.053	1.388	2.831
Observations	164,966	164,966	164,966	164,966	164,966
Panel B. Stratification by hospitalization history					
(1) No hospitalization history					
Kessner adequacy	-0.0346**	-0.0215	0.0103	-1.7981	-3.6051
	(0.0111)	(0.0132)	(0.0082)	(1.3949)	(2.1777)
Mean of dep var	0.131	0.118	0.048	1.391	2.996
Observations	317,750	317,750	317,750	317,750	317,750
(2) With hospitalization history					
Kessner adequacy	-0.0281	-0.0289	-0.0023	-5.4332	-2.1234
	(0.0489)	(0.0523)	(0.0419)	(6.2973)	(9.9820)
Mean of dep var	0.123	0.118	0.064	1.789	4.136
Observations	17,890	17,890	17,890	17,890	17,890
Panel C. Weak first-stage region analysis					
(1) Excluding East Regions (IV Estimation)					
Kessner adequacy	-0.0341**	-0.0221	0.0114	-2.0175	-3.1283
	(0.0106)	(0.0125)	(0.0081)	(1.3562)	(2.1754)
Mean of dep var	0.132	0.118	0.049	1.408	3.048
Observations	328,126	328,126	328,126	328,126	328,126
(2) East Regions (Reduced Form)					
Exposure fraction	0.0084	-0.0059	0.0219	-1.2514	4.5016
	(0.0175)	(0.0213)	(0.0150)	(2.6206)	(4.2314)
Mean of dep var	0.087	0.115	0.051	1.590	3.445
Observations	7,548	7,548	7,548	7,548	7,548
Risk factors	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes
Birth month FE	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Effects of SARS Exposure on Diagnosis and Delivery Mode

	(1)	(2)	(3)
	Anemia	Hypertensive Disorders of Pregnancy	Cesarean Section
Exposure fraction	-0.0081***	-0.0009	-0.0201***
	(0.0016)	(0.0010)	(0.0050)
Risk factors	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes
Birth month FE	Yes	Yes	Yes
Mean of dep var	0.026	0.009	0.344
Observations	332,746	332,746	332,746

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: We use ICD-9 codes to identify anemia and hypertensive disorders of pregnancy. See Appendix Table A6.1 Panel C for references.

Online Appendix

When More Prenatal Care Looks Like Less:

Evidence on Selection and Detection from the SARS Epidemic

Chien-Tzu Cheng
Shin-Yi Chou
Hsien-Ming Lien

November 2026

A1 Epidemiological Context: The 2003 SARS Outbreak in Taiwan

The first human case was retrospectively identified on November 16, 2002, in the southern Chinese province of Guangdong. The virus spread internationally after a medical doctor who had treated patients in Guangdong traveled to Hong Kong; guests at the same hotel subsequently carried the virus to other countries. On March 12, 2003, the WHO issued its first global alert, describing outbreaks of the yet-unnamed respiratory disease in both Hong Kong and Vietnam (Knobler et al., 2004).

Despite extensive business interactions between Taiwan, Hong Kong, and China, SARS transmission remained sporadic in Taiwan during March and early April 2003. The situation escalated on April 22, when a nosocomial cluster at Heping Hospital in Taipei—ignited by an unrecognized SCoV-carrier—triggered facility quarantine and a nationwide public health emergency, eventually linked to 26 deaths. Health authorities responded by tightening surveillance, strengthening infection control measures, and launching a mass education campaign. Figure A1.1 shows the geographic distribution of probable cases as of May 22, 2003, with the majority concentrated in Taipei City and Taipei County (now New Taipei City).

Figure A1.2 plots the average number of prenatal care visits per birth from 2001 to 2006. Prior to the SARS outbreak, visit counts were stable at approximately 9.3–9.5 visits per birth. Following the emergence of the first confirmed case in Taiwan on March 14, 2003, visits declined sharply, falling to roughly 7.3 by mid-2003 and remaining near that trough through 2004. The recovery was gradual: visit counts began rising slowly only in mid-2005 and did not return to pre-epidemic levels until after 2006, underscoring that the behavioral response to SARS—fear of visiting healthcare facilities—persisted well beyond the WHO's declaration of worldwide containment on July 5, 2003.

Figure A1.3 plots the total number of prenatal visits by birth week separately for the East region and all other regions of Taiwan. Prior to the SARS outbreak, the two groups followed nearly identical trajectories, with prenatal visit counts stable at approximately 9.3–9.5 visits across both regions, providing reassurance that the East region serves as a credible comparison group. Following the onset of the epidemic in March 2003, a sharp divergence emerges: prenatal visits in other regions declined precipitously, falling by more than 20% to approximately 7.3 visits, while the East region experienced virtually no disruption, with visit counts remaining stable throughout the outbreak period. This regional contrast is consistent with the geographic concentration of SARS cases documented in Figure A1.1, where confirmed cases were overwhelmingly concentrated in Taipei City and New Taipei City, with the East region largely spared from the epidemic.

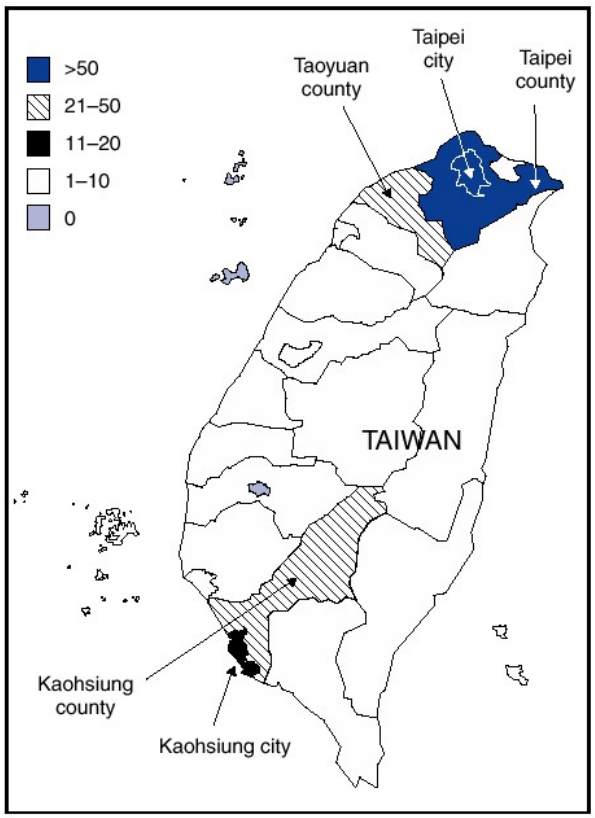


Figure A1.1: Geographic Distribution of Probable Case of SARS in Taiwan

Notes: N= 483. As of May 22.
Source: *Centers for Disease Control and Prevention (CDC)* (2003) Accessed from here.

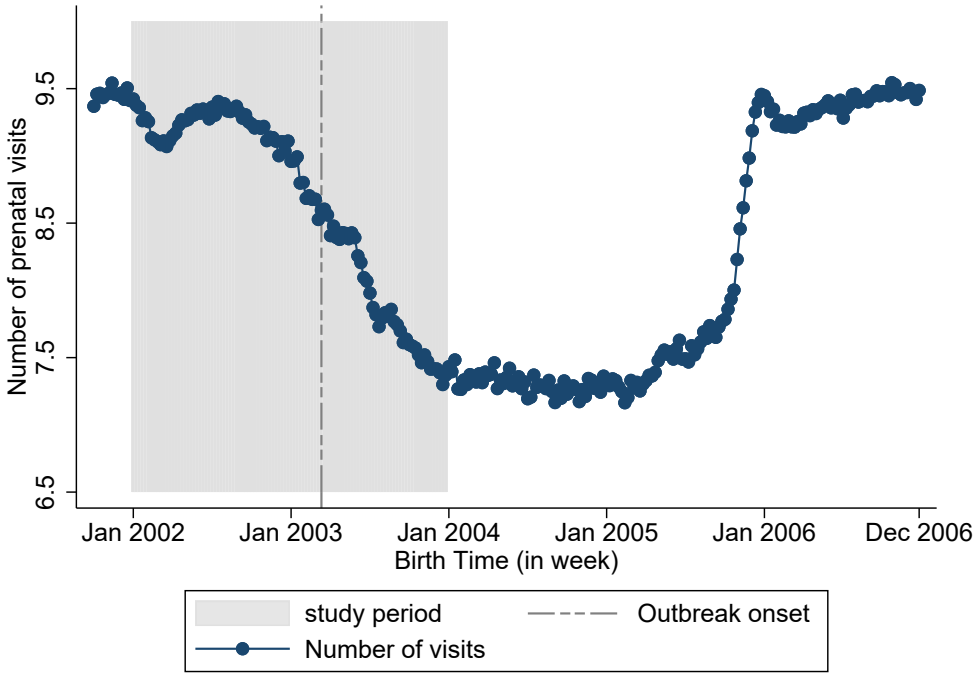


Figure A1.2: Numbers of Total Visits Over Time

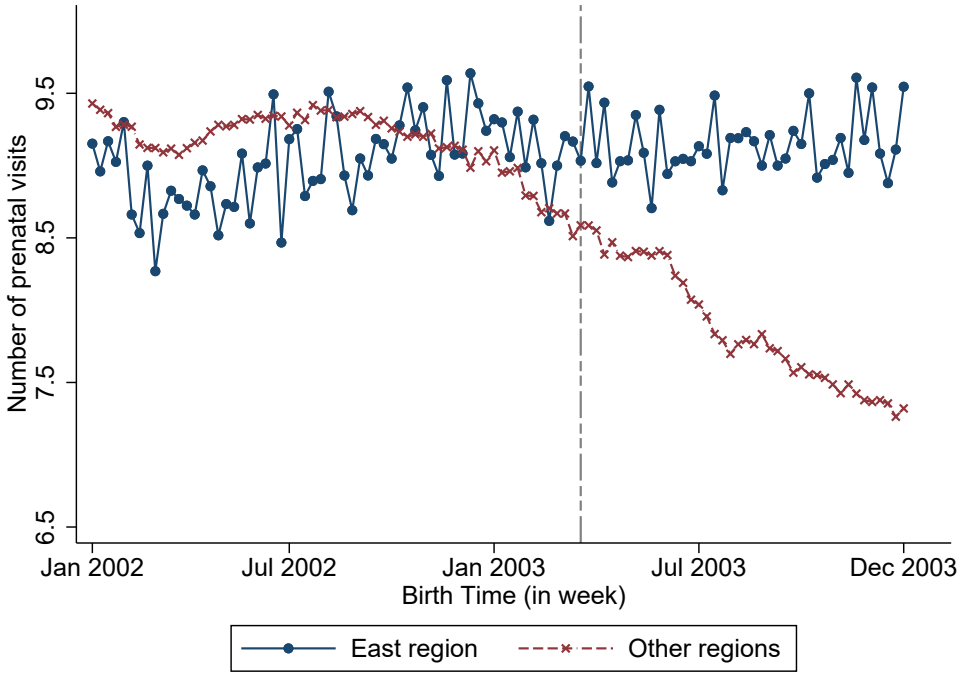


Figure A1.3: Numbers of Total Visits Over Time by Region

A2 The NHI Prenatal Care Schedule: Visit Content and Adherence

Prenatal Care Schedule

This appendix provides the full visit-by-visit description of Taiwan's NHI prenatal care schedule summarized in Section 2.3 of the main text. Under the schedule in effect during our study period, NHI covered ten prenatal checkups distributed across pregnancy. Importantly for our analysis, the schedule combined a common set of assessments performed routinely at every checkup with additional laboratory and screening services scheduled at particular gestational stages. This structure allowed repeated visits to generate a longitudinal clinical record while making attendance at certain visits particularly important for receiving services not included in routine monitoring. Appendix Figure A2.1 reproduces the contemporaneous schedule from the government-issued Maternal Health Booklet.

A common set of assessments, labeled the "routine prenatal test," was performed at each scheduled checkup. These included questions about pregnancy-related symptoms or discomfort; physical assessments of maternal weight and blood pressure, fetal heartbeat and position, and edema; and urine testing for protein and glucose. Thus, blood pressure and urine protein were monitored repeatedly throughout pregnancy as part of the routine content of prenatal care. By contrast, blood testing for hemoglobin and hematocrit was included in the laboratory panel at the initial visit but was not listed as part of the routine test performed at subsequent checkups. This distinction is central to the detection analysis in Section 7: conditions identifiable through routine measurements could potentially be detected whenever a scheduled visit occurred, whereas detecting anemia that developed after the initial blood panel required additional laboratory testing beyond the standard content of later visits.

First trimester (before week 17). The *first visit*, recommended before week 12, was the most comprehensive. Following confirmation of a fetal heartbeat, the schedule called for assessment of family and obstetric history, past medical history, pregnancy history, and pregnancy-related symptoms and behaviors; physical examination including weight, height, blood pressure, thyroid and breast examinations, and pelvic and abdominal examinations; and an extensive laboratory panel. The blood tests included white and red blood cell counts, platelet count, hematocrit, hemoglobin, mean corpuscular volume, blood type and Rh factor, hepatitis B surface and e antigens, syphilis testing, rubella IgG, and HIV testing, together with routine urine testing. The *second visit*, recommended at week 16, included the routine prenatal test together with guidance on preventing premature birth.

Second trimester (weeks 17–28). The *third visit*, recommended at week 20, was distinguished from surrounding routine visits by the inclusion of a scheduled ultrasound examination, in addition to the routine prenatal test and guidance on premature-birth prevention. The Maternal Health Booklet specifically provided that women unable to undergo the ultrasound at this visit because of special circumstances could receive it instead in the third trimester. The *fourth visit*, recommended at week 28, consisted of the routine prenatal test. Thus, the week-20 visit represented a gestationally

timed opportunity to obtain fetal imaging that was not part of the routine content of every prenatal checkup.

Third trimester (after week 29). The remaining six checkups became progressively more frequent as delivery approached and combined continued routine surveillance with additional tests at selected visits. The *fifth visit*, recommended at week 32, included the routine prenatal test together with syphilis screening (VDRL or RPR); an additional HIV test was recommended for women identified as being at risk of HIV infection. The *sixth visit*, at week 34, consisted of the routine prenatal test. The *seventh visit*, at week 36, combined the routine prenatal test with subsidized maternal Group B Streptococcus screening. The final three visits, recommended at weeks 38, 39, and 40, respectively, continued the routine prenatal assessment as delivery approached.

The schedule therefore combined repeated surveillance with gestationally timed diagnostic and screening services. Blood pressure and urine protein and glucose were assessed as part of every routine prenatal checkup, generating repeated opportunities to identify evolving abnormalities. Other clinical information was obtained through services tied to particular stages of pregnancy, including the comprehensive laboratory panel at the initial visit, the scheduled ultrasound at week 20, repeat syphilis testing at week 32, and Group B Streptococcus screening at week 36. Missing prenatal visits could therefore matter in two ways: by interrupting the accumulation of information from repeated routine monitoring and by reducing access to services scheduled at particular gestational stages. Continued attendance consequently determined not only the quantity of prenatal care a woman received, but also the clinical information available to providers as pregnancy progressed. The full NHI prenatal care timeline is reproduced in Appendix Figure A2.1.

Trimester-Specific Adequacy Before and After SARS

Figure A2.2 plots the probability of achieving trimester-specific prenatal care adequacy against birth timing, covering 2002–2006—an extended window relative to the 2002–2003 study period used in the main analysis. The figure shows three series, corresponding to the completion of all NHI-subsidized checkups within each trimester: first trimester (two visits), second trimester (two visits), and third trimester (six visits).

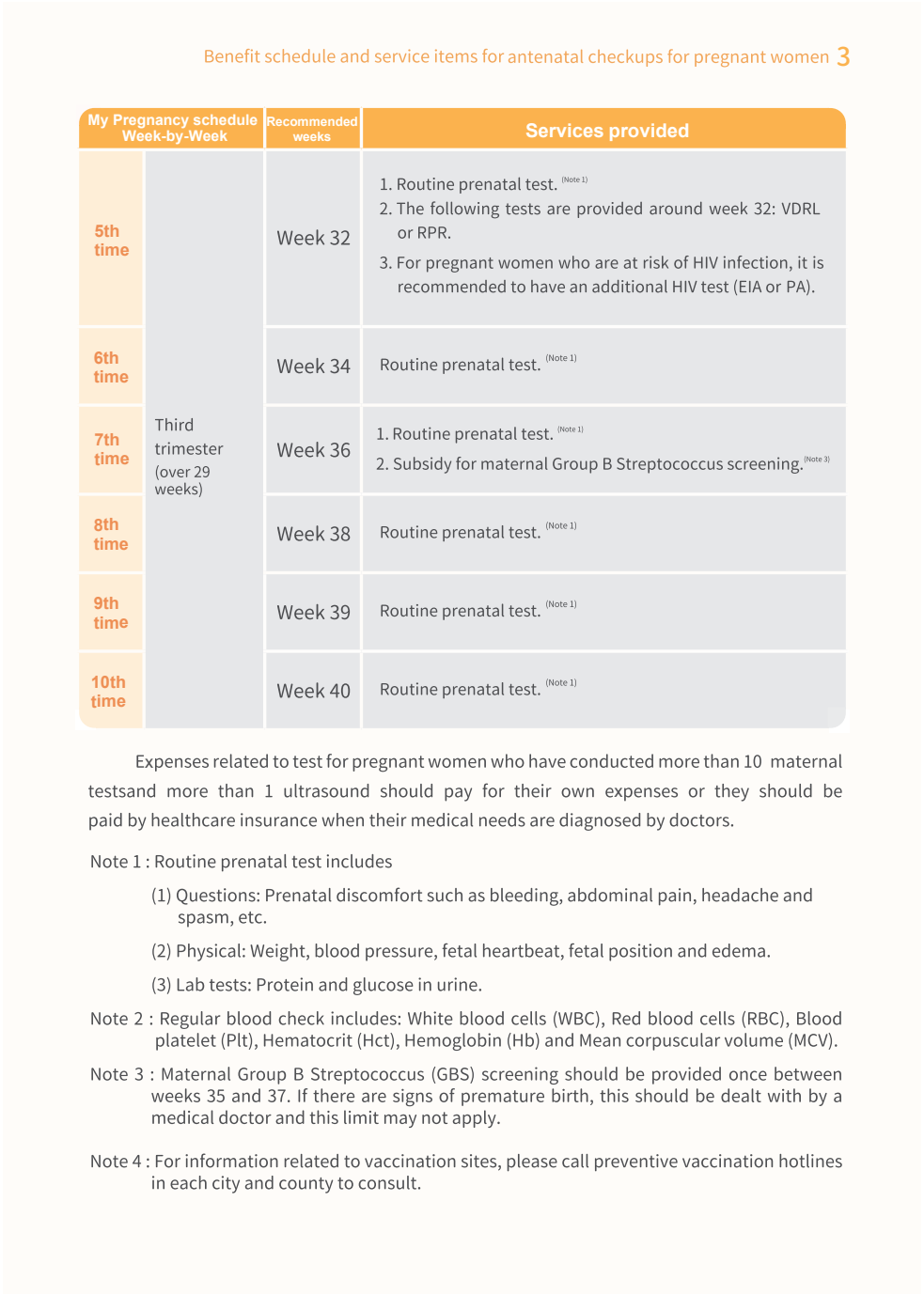
All three series are stable prior to the outbreak, then decline following its onset, with the second-trimester series exhibiting the steepest decline. Following the end of the study period, all three series recover gradually, returning to or exceeding pre-SARS levels by 2005–2006—mirroring the protracted recovery in total prenatal visits shown in Figure A1.2.

2

My Pregnancy schedule Week-by-Week for prenatal test and services provided



My Pregnancy schedule Week-by-Week		Recommended weeks	Services provided
1st time	First trimester (Less than 17 weeks)	Prior to week 12	1. After confirming that there is a fetal heartbeat, the following inspections must be included: (1) Questions: Family history, expectant mother's past medical history, pregnancy history and any discomfort with the fetus,addicted habits. (2) Physical: Weight, height, blood pressure, thyroid, breast exam,pelvis exam, chest and abdomen exams. (3) Lab testing: Blood routine (WBC, RBC, Plt, Hct , Hb, MCV),blood type, Rh factor, HBsAg and HBeAg (Thosewho cannot go through these exams this time due tospecial circumstances can do so in the 5th prenataltest instead),VDRL (syphilis test) , RPR(Rapid plasma reagin), Rubella IgG (rubellanti-body), AIDS (EIA or PA) and urine routine. (Note 2) 2. Routine renatal test. (Note 1) Note: Pregnant women who tested negative for the Rubella antibody should be vaccinated after giving birth, this vaccine is government-funded. (Note 4)
2nd time		Week 16	1 . Routine prenatal test. (Note 1) 2 . Premature birth prevention guide.
3rd time	Second trimester (over 17 weeks but less than 29 weeks)	Week 20	1 . Routine prenatal test. (Note 1) 2 . Ultrasound (Those who cannot go through ultrasound at this time due to special circumstances can do so in the third trimester instead). 3 . Premature birth prevention guide.
4th time		Week 28	Routine prenatal test. (Note 1)



Benefit schedule and service items for antenatal checkups for pregnant women 3

My Pregnancy schedule Week-by-Week		Recommended weeks	Services provided
5th time	Third trimester (over 29 weeks)	Week 32	1. Routine prenatal test. (Note 1) 2. The following tests are provided around week 32: VDRL or RPR. 3. For pregnant women who are at risk of HIV infection, it is recommended to have an additional HIV test (EIA or PA).
6th time		Week 34	Routine prenatal test. (Note 1)
7th time		Week 36	1. Routine prenatal test. (Note 1) 2. Subsidy for maternal Group B Streptococcus screening. (Note 3)
8th time		Week 38	Routine prenatal test. (Note 1)
9th time		Week 39	Routine prenatal test. (Note 1)
10th time		Week 40	Routine prenatal test. (Note 1)

Expenses related to test for pregnant women who have conducted more than 10 maternal testsand more than 1 ultrasound should pay for their own expenses or they should be paid by healthcare insurance when their medical needs are diagnosed by doctors.

Note 1 : Routine prenatal test includes

(1) Questions: Prenatal discomfort such as bleeding, abdominal pain, headache and spasm, etc.

(2) Physical: Weight, blood pressure, fetal heartbeat, fetal position and edema.

(3) Lab tests: Protein and glucose in urine.

Note 2 : Regular blood check includes: White blood cells (WBC), Red blood cells (RBC), Blood platelet (Plt), Hematocrit (Hct), Hemoglobin (Hb) and Mean corpuscular volume (MCV).

Note 3 : Maternal Group B Streptococcus (GBS) screening should be provided once between weeks 35 and 37. If there are signs of premature birth, this should be dealt with by a medical doctor and this limit may not apply.

Note 4 : For information related to vaccination sites, please call preventive vaccination hotlines in each city and county to consult.

Figure A2.1: Prenatal Care Timeline from Maternal Health Booklet

Source: *Centers for Disease Control and Prevention (CDC)* (2003)

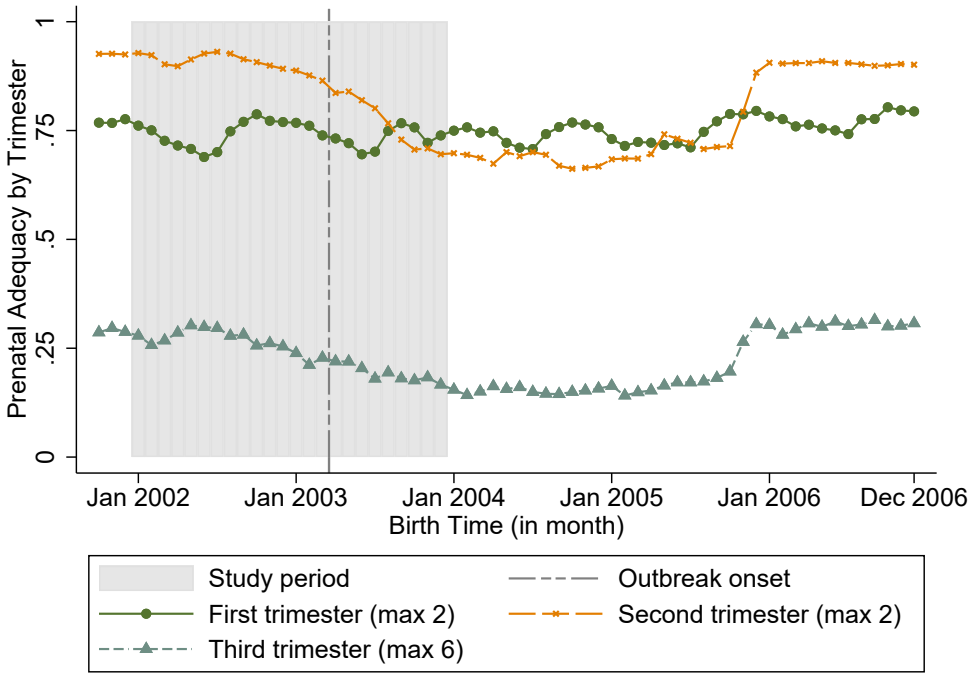


Figure A2.2: Probability of Achieving Trimester-Specific Prenatal Care Adequacy Relative to Birth Timing

Notes: We define trimester adequacy as the completion of all NHI subsidized check-ups within a given period: two visits in both the first and second trimesters, and six in the third. We note that the baseline probability of adequacy is naturally lower in the third trimester due to biological variation in gestational length.

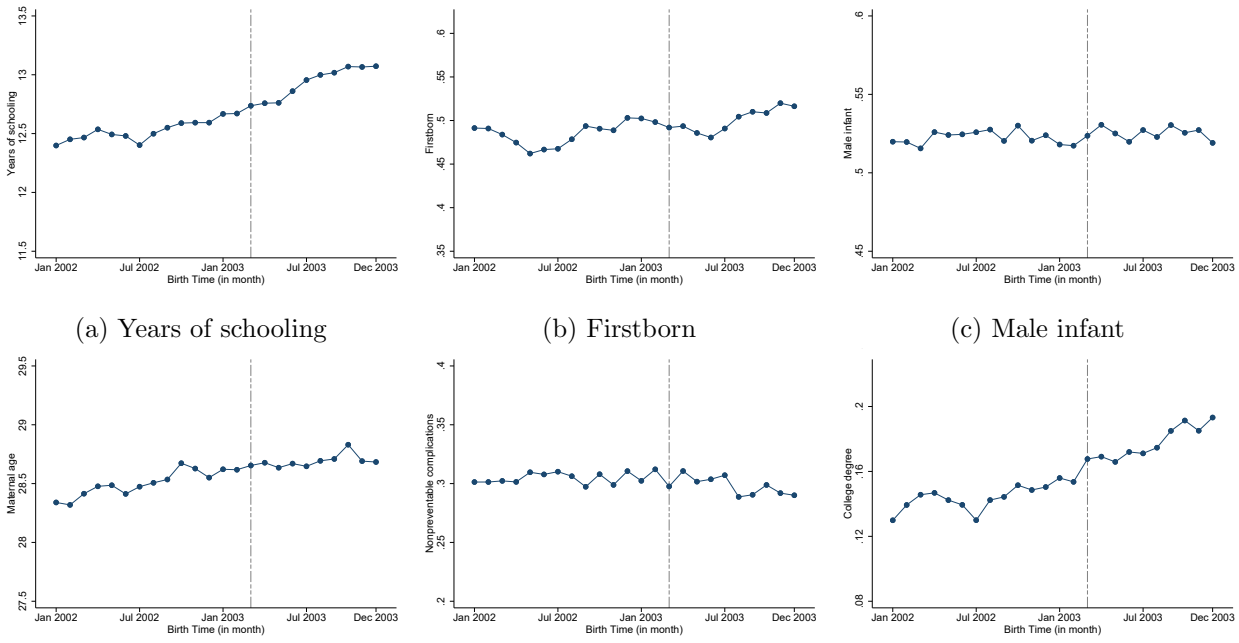
A3 Addressing the Education Gradient in SARS Exposure

Table 1 and the balance tests in Section 5.2 Table 3 Panel A document one systematic difference between the exposed and unexposed groups: women who delivered after the onset of SARS have slightly higher educational attainment, averaging 12.95 versus 12.53 years of schooling, with a roughly 3-percentage-point higher share of college graduates. While statistically significant given our large sample, this gradient is economically modest and, we argue, reflects secular trends in educational attainment rather than endogenous selection into SARS exposure.

Figure A3.1 provides the primary evidence for this interpretation. Panel (a) plots mean years of schooling by birth month across the full 2002–2003 sample period. The series displays a smooth, continuous upward trend with no detectable discontinuity at the onset of the SARS outbreak in March 2003. Panel (f) shows the same pattern for the college-degree share specifically. If the education gradient were driven by selective fertility timing—for instance, if highly educated women disproportionately postponed conception in response to SARS—we would expect a discrete shift in educational attainment around the onset of the epidemic. Instead, the series passes through the outbreak onset without interruption, consistent with the gradient reflecting a pre-existing secular trend that happens to coincide with the later delivery dates of the exposed group. The remaining panels confirm that other maternal and birth characteristics—firstborn status, infant sex, maternal age, and nonpreventable complication rates—show no systematic trend or discontinuity across the study period, reinforcing that the education gradient is the exception rather than the rule.

Nonetheless, the gradient raises a legitimate concern: to the extent that college-educated mothers are disproportionately represented in the exposed group, and if their behavioral or health responses to SARS differ systematically from those of less-educated mothers, our main estimates could be distorted. We assess this directly through a series of robustness exercises reported in Table A3.1. We randomly drop 1%, 5%, and 10% of college-educated mothers from the exposed group and re-estimate all specifications—OLS, reduced-form, and IV—across the full set of maternal and infant outcomes.

The results are strikingly stable. Across all three subsamples and all three estimators, point estimates are nearly identical to those reported in the main analysis. The pattern of opposing selection bias across outcome types is fully preserved: OLS overstates the benefits of prenatal care for infant outcomes and understates them for maternal complications, while IV estimates concentrate the causal effect on preventable maternal complications with no statistically significant effects on infant outcomes. The stability of these estimates across subsamples that progressively reduce the college-educated share of the exposed group provides additional evidence that the education gradient does not meaningfully distort our conclusions. Our main findings are not an artifact of compositional differences between the exposed and unexposed groups.



(a) Years of schooling (b) Firstborn (c) Male infant

(d) Maternal age (e) Nonpreventable complications (f) College degree

Figure A3.1: Trends in Maternal and Birth Characteristics by Birth Time

Table A3.1: Estimated Effects of Prenatal Care, Randomly Dropping Women With College Degrees

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Preventable Complications		SGA		LBW		Neonatal Mortality		Infant Mortality	
Panel A. OLS estimation										
Kessner adequacy (drop 1%)	0.0114***	0.0096***	-0.0066***	-0.0063***	-0.0447***	-0.0438***	-0.7930***	-0.7820***	-1.2565***	-1.2365***
	(0.0011)	(0.0010)	(0.0013)	(0.0013)	(0.0010)	(0.0010)	(0.1602)	(0.1600)	(0.2380)	(0.2385)
Kessner adequacy (drop 5%)	0.0114***	0.0095***	-0.0067***	-0.0063***	-0.0447***	-0.0439***	-0.8036***	-0.7920***	-1.2745***	-1.2537***
	(0.0011)	(0.0010)	(0.0013)	(0.0013)	(0.0010)	(0.0010)	(0.1610)	(0.1608)	(0.2383)	(0.2387)
Kessner adequacy (drop 10%)	0.0114***	0.0096***	-0.0066***	-0.0063***	-0.0447***	-0.0438***	-0.8048***	-0.7937***	-1.2820***	-1.2618***
	(0.0011)	(0.0010)	(0.0013)	(0.0013)	(0.0010)	(0.0010)	(0.1613)	(0.1611)	(0.2388)	(0.2392)
Panel B. Reduced-form estimation										
Exposure fraction (drop 1%)	0.0117**	0.0095**	0.0061	0.0062	-0.0033	-0.0027	0.5047	0.5300	0.9457	0.9771
	(0.0041)	(0.0030)	(0.0036)	(0.0035)	(0.0023)	(0.0023)	(0.3848)	(0.3858)	(0.6153)	(0.6152)
Exposure fraction (drop 5%)	0.0121**	0.0100***	0.0060	0.0060	-0.0035	-0.0029	0.5298	0.5559	0.9098	0.9423
	(0.0041)	(0.0030)	(0.0036)	(0.0035)	(0.0023)	(0.0023)	(0.3859)	(0.3868)	(0.6164)	(0.6163)
Exposure fraction (drop 10%)	0.0121**	0.0098**	0.0063	0.0063	-0.0033	-0.0026	0.4694	0.4959	0.8867	0.9197
	(0.0041)	(0.0030)	(0.0036)	(0.0036)	(0.0023)	(0.0023)	(0.3869)	(0.3879)	(0.6199)	(0.6198)
Panel C. IV estimation										
Kessner adequacy (drop 1%)	-0.0416**	-0.0338**	-0.0217	-0.0218	0.0117	0.0096	-1.7925	-1.8793	-3.3591	-3.4651
	(0.0147)	(0.0107)	(0.0127)	(0.0125)	(0.0081)	(0.0081)	(1.3659)	(1.3670)	(2.1891)	(2.1855)
Kessner adequacy (drop 5%)	-0.0430**	-0.0353**	-0.0211	-0.0213	0.0124	0.0102	-1.8757	-1.9650	-3.2208	-3.3308
	(0.0147)	(0.0107)	(0.0127)	(0.0126)	(0.0081)	(0.0081)	(1.3651)	(1.3662)	(2.1856)	(2.1821)
Kessner adequacy (drop 10%)	-0.0428**	-0.0346**	-0.0222	-0.0224	0.0116	0.0094	-1.6655	-1.7565	-3.1459	-3.2575
	(0.0147)	(0.0107)	(0.0127)	(0.0126)	(0.0081)	(0.0082)	(1.3719)	(1.3729)	(2.2026)	(2.1989)
Risk factors	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A4 First Stage Estimation and Robustness Checks

Alternative Measures for Prenatal Care Utilization

We replace our main prenatal measure, the Kessner adequacy index, with two alternative measures. First, we use a continuous count of prenatal visits that captures the intensity of care use. Second, we create an indicator for receiving at least 10 visits, which reflects whether a woman has met the recommended minimum threshold. Table A4.1 reports a significant first stage for both alternative measures. Columns (3) and (4) show that a 100% increase in SARS exposure fraction reduces total prenatal care visits by 1.4 visits (or 16%); columns (5) and (6) show that a 100% increase in SARS exposure fraction reduces the likelihood of receiving at least 10 visits by 26 percentage points (or 53%). SARS exposure significantly predicts prenatal care utilization estimated by these two measures.

Table A4.2 Panel A reports the main results using the number of total visits as the measure, Panel B using an indicator for at least 10 visits as the measure. Overall, we see a consistent pattern: the OLS estimation underestimates the effects on maternal outcomes and overestimates the effects on infant birth outcomes. Panels A.1 and B.1 show that the OLS specification yields positive coefficients for maternal outcomes and negative coefficients for infant outcomes. Panels A.2 and B.2 show that the reduced-form addresses the bias and finds significant positive effects of exposure on maternal outcomes and negligible effects on infant outcomes. The IV estimation presents similar patterns. Panel A.3 shows that an additional prenatal care visit significantly reduces the likelihood of preventable complications by 0.7 percentage points (or 5%). Panel B.3 shows that following the medical guideline to receive at least 10 visits significantly reduces the likelihood of preventable complications by approximately 4 percentage points (28%). Meanwhile, the IV estimation reports no detectable impact on infant birth outcomes. The consistency estimates across prenatal care measures indicates that the findings are not driven by the particular measure of prenatal care adequacy.

Controlling for Conception Month Seasonality

In our primary specifications, we include birth-month fixed effects and a birth-month linear time trend to control for selection in seasonality, as birth month is found to systematically correlate with maternal characteristics (Buckles and Hungerman, 2013). It also helps to address common shocks that affect births in the same month. For example, major holidays can affect the timing of elective deliveries. However, other studies suggest that seasonality in birth outcomes is largely driven by systematic variation in maternal characteristics among women at the time of conception (Dehejia and Lleras-Muney, 2004; Currie and Schwandt, 2013). Since gestational length is endogenous and the exact delivery date is not perfectly predictable, controlling for the conception month may more accurately capture the initial health environment and maternal sorting.

To address this, we re-estimate our baseline model by replacing the birth-month-based controls with conception-month fixed effects and a conception-month linear time trend. Table A4.3 presents

Table A4.1: First Stage Estimation Using Three Prenatal Care Measures and East Region Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full Sample		Full Sample		Full Sample		East Region	
	Kessner adequacy		Total visits		At least 10 visits		Kessner adequacy	
Exposure fraction	-0.282***	-0.282***	-1.411***	-1.414***	-0.262***	-0.263***	-0.047	-0.048
	(0.006)	(0.006)	(0.028)	(0.028)	(0.006)	(0.006)	(0.031)	(0.031)
Risk factors	No	Yes	No	Yes	No	Yes	No	Yes
Mean of dep var	0.637	0.637	8.677	8.677	0.493	0.493	0.687	0.687
Observations	335,674	335,674	335,674	335,674	335,674	335,674	7,548	7,548
Instrument F	2176	2172	2615	2610	1805	1818	2	2

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

these results. The point estimates across all specifications remain nearly identical to our baseline results in Table 4, suggesting that our findings are robust to alternative specifications of seasonality.

Nonlinear Probability Models: IV Probit

Given that our dependent variables are binary, we examine whether our linear probability model (LPM) estimates are sensitive to the nonlinear structure of the underlying probability distribution.[25] We adopt an IV probit model to improve estimator precision in the context of rare events, such as neonatal and infant mortality.

Table A4.4 reports the estimated coefficients from the IV probit specifications. Consistent with our primary findings, we document significant negative effects of adequate prenatal care on preventable complications, while the impacts on infant outcomes remain statistically insignificant. The qualitative conclusions regarding the lack of impact on infant mortality remain unchanged, and the computed marginal effects are highly consistent with the linear probability model baseline.

[25] In the main specification, we multiply the outcomes of neonatal and infant mortality by 1,000 to estimate the effects of prenatal care on mortality rates per 1,000 births. Here, in the IV probit model, outcomes of neonatal and infant mortality are binary, indicating death within 28 days and a year, respectively.

Table A4.2: Additional Specification Using Alternative Prenatal Care Utilization Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Preventable Complications		SGA		LBW		Neonatal Mortality		Infant Mortality	
Panel A. Prenatal care measure: the number of total visits										
Panel A.1 OLS estimation										
Total Visit	0.0030***	0.0024***	-0.0011***	-0.0010***	-0.0141***	-0.0138***	-0.3737***	-0.3698***	-0.4843***	-0.4777***
	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0419)	(0.0418)	(0.0602)	(0.0602)
Panel A.2 Reduced-form estimation										
Exposure fraction	0.0118**	0.0097**	0.0060	0.0060	-0.0034	-0.0028	0.5212	0.5468	0.9554	0.9870
	(0.0041)	(0.0030)	(0.0036)	(0.0035)	(0.0023)	(0.0023)	(0.3849)	(0.3859)	(0.6148)	(0.6147)
Panel A.3 IV estimation										
Total Visit	-0.0084**	-0.0069**	-0.0043	-0.0043	0.0023	0.0019	-0.3697	-0.3871	-0.6744	-0.6954
	(0.0029)	(0.0021)	(0.0025)	(0.0025)	(0.0016)	(0.0016)	(0.2723)	(0.2724)	(0.4361)	(0.4351)
Panel B. Prenatal care measure: An indicator for visits ≥ 10										
Panel B.1 OLS estimation										
10 prenatal visits	0.0191***	0.0136***	-0.0026*	-0.0024*	-0.0562***	-0.0550***	-1.0806***	-1.0594***	-1.7400***	-1.7018***
	(0.0011)	(0.0010)	(0.0012)	(0.0012)	(0.0009)	(0.0009)	(0.1418)	(0.1411)	(0.2172)	(0.2178)
Panel B.2 Reduced-form estimation										
Exposure fraction	0.0118**	0.0097**	0.0060	0.0060	-0.0034	-0.0028	0.5212	0.5468	0.9554	0.9870
	(0.0041)	(0.0030)	(0.0036)	(0.0035)	(0.0023)	(0.0023)	(0.3849)	(0.3859)	(0.6148)	(0.6147)
Panel B.3 IV estimation										
10 prenatal visits	-0.0455**	-0.0369**	-0.0232	-0.0234	0.0124	0.0101	-1.9932	-2.0832	-3.6359	-3.7423
	(0.0159)	(0.0115)	(0.0137)	(0.0135)	(0.0086)	(0.0087)	(1.4690)	(1.4663)	(2.3532)	(2.3439)
Risk factors	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	0.131	0.131	0.118	0.118	0.049	0.049	1.412	1.412	3.057	3.057
Observations	335,674	335,674	335,674	335,674	335,674	335,674	335,674	335,674	335,674	335,674

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The reduced-form estimates, which regress the outcomes directly on SARS exposure, are identical across specifications regardless of the first-stage definition. They are reported here for completeness and convenience.

Table A4.3: Additional Specification Replacing Birth Month With Conception Month as Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Preventable Complications		SGA		LBW		Neonatal Mortality		Infant Mortality	
Panel A. OLS estimation										
Kessner adequacy	0.0114***	0.0095***	-0.0066***	-0.0063***	-0.0446***	-0.0438***	-0.7996***	-0.7881***	-1.2632***	-1.2427***
	(0.0011)	(0.0010)	(0.0013)	(0.0013)	(0.0010)	(0.0010)	(0.1605)	(0.1603)	(0.2370)	(0.2375)
Panel B. Reduced-form estimation										
Exposure fraction	0.0117**	0.0093**	0.0061	0.0062	-0.0031	-0.0025	0.5945	0.6202	1.0589	1.0912
	(0.0041)	(0.0030)	(0.0036)	(0.0035)	(0.0023)	(0.0024)	(0.3837)	(0.3844)	(0.6169)	(0.6166)
Panel C. IV estimation										
Kessner adequacy	-0.0419**	-0.0333**	-0.0219	-0.0220	0.0110	0.0088	-2.1250	-2.2131	-3.7853	-3.8939
	(0.0147)	(0.0107)	(0.0128)	(0.0127)	(0.0084)	(0.0084)	(1.3724)	(1.3727)	(2.2138)	(2.2094)
Risk factors	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Conception month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	0.131	0.131	0.118	0.118	0.049	0.049	1.412	1.412	3.057	3.057
Observations	335,674	335,674	335,674	335,674	335,674	335,674	335,674	335,674	335,674	335,674

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: In this alternative specification, we replace birth month fixed effects and linear calendar birth month time trend effects with conception fixed effects and linear calendar conception month time trend effects.

Table A4.4: Additional Specification Using IV Probit Model

	(1) Prev. Complications	(2) SGA	(3) LBW	(4) Neonatal Mortality	(5) Infant Mortality
Kessner adequacy	-0.342***	-0.121	0.080	-0.426	-0.357
	(0.090)	(0.066)	(0.081)	(0.280)	(0.224)
Risk factors	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes
Birth month FE	Yes	Yes	Yes	Yes	Yes
Mean of dep var	0.185	0.118	0.049	0.002	0.003
Observations	236,851	335,618	334,911	254,320	298,515

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A5 Characterizing the Treated Margin: Who Are the Compliers?

A standard concern with instrumental variables estimation is that the identified effect may speak only to a narrow subpopulation of compliers—women whose prenatal care utilization was shifted by the instrument—rather than to the broader population of interest. If the SARS shock disproportionately disrupted care among a particular demographic subgroup, our LATE would have limited policy relevance beyond that group. We assess this concern by examining heterogeneity in the first-stage response to SARS within a Generalized Random Forest (GRF) framework, following Wager and Athey (2018) and Athey et al. (2019).

We estimate Conditional Average Treatment Effects (CATEs) for prenatal care adequacy and partition observations into quintiles according to their predicted responsiveness to the shock. The covariates employed mirror the birth-specific control set used throughout the analysis, comprising maternal age, education, infant sex, firstborn status, and nonpreventable complications. Figure A5.1 plots the average realized treatment effect within each quintile, estimated independently using the Augmented Inverse Probability Weighting (AIPW) estimator.

Two features of Figure A5.1 are worth noting. First, the AIPW estimates are non-monotonic across quintiles—they do not rise systematically from the first to the fifth quintile as they would if the GRF's predicted ranking reliably sorted women by their true responsiveness to the shock. Second, the confidence intervals across quintiles overlap substantially throughout, and no pairwise difference approaches statistical significance. This pattern is meaningful precisely because the quintile ranking is constructed from the GRF's predicted responsiveness scores: if observable covariates reliably sorted women by true responsiveness, the AIPW estimates should display a monotonic gradient. Their failure to do so indicates that observable characteristics have limited power to identify who responded more or less to the SARS shock.

Figure A5.2 sheds light on what the GRF is using to generate its predicted ranking. Among all covariates, maternal education exhibits the most pronounced gradient across quintiles, rising monotonically from 11.5 years of schooling in the first quintile to 14.8 years in the fifth. By contrast, maternal age, nonpreventable complications, firstborn status, and infant sex show no systematic pattern across quintiles. The GRF is thus primarily using education to predict responsiveness—yet as Figure A5.1 demonstrates, this education-based ranking does not translate into meaningful differences in realized treatment effects. The education gradient visible in Figure A5.2 does not reflect genuine behavioral heterogeneity in response to SARS; as established in Section 3.4 and Appendix A3, it reflects the secular upward trend in educational attainment that coincides with later delivery dates in our sample.

Taken together, the non-monotonic AIPW estimates and the absence of any gradient in age, parity, and clinical risk across quintiles indicate that the SARS epidemic generated a broad and relatively uniform first-stage shock. This has an important implication for external validity. Most existing IV studies in the prenatal care literature identify effects for narrow complier populations concentrated among low-income or low-access subgroups, where the instrument—such as a Medicaid expansion or a bus strike—disproportionately shifts behavior among the most disadvantaged

women. In our setting, the SARS shock does not appear to have been concentrated among a narrowly defined subgroup based on the observable characteristics available in our data. This suggests a broader complier population than in many prenatal-care IV studies, although unobserved heterogeneity in compliance may remain. While we cannot formally claim that our LATE equals the Average Treatment Effect—unobservable characteristics such as baseline health consciousness or individual fear responses to SARS could still jointly predict compliance and potential outcomes in ways the GRF cannot detect—the absence of systematic heterogeneity across a rich set of observable characteristics provides some reassurance about the external validity of our estimates. Our identifying variation therefore appears less concentrated among specific observable subpopulations than in many prior prenatal-care studies, particularly those whose instruments operate primarily through changes in access among disadvantaged groups.

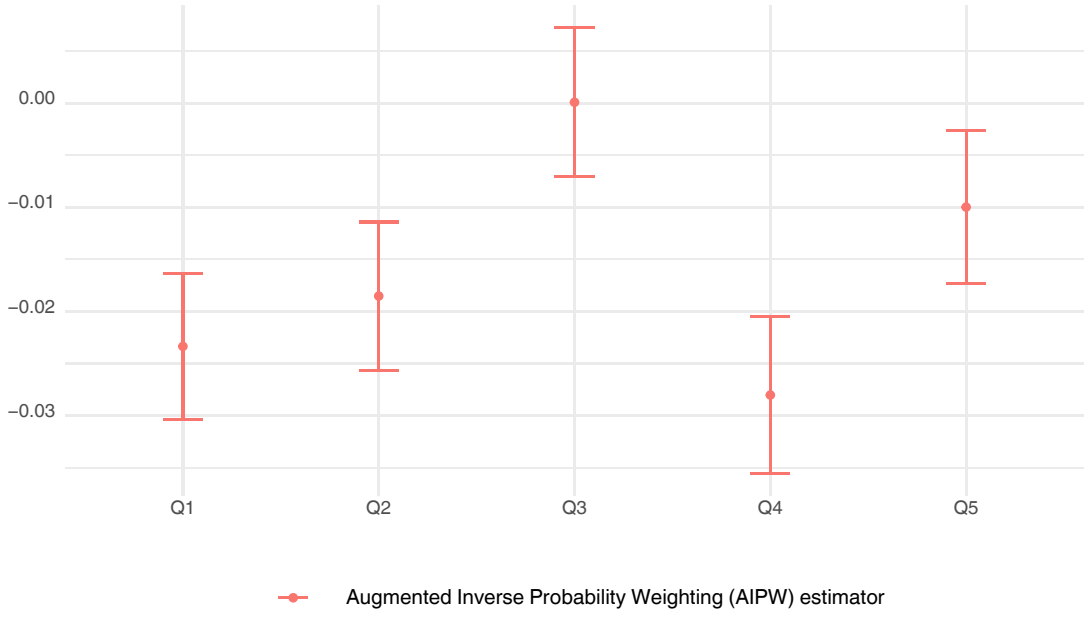


Figure A5.1: Conditional Average Treatment Effect (CATE) Within Each Ranking (as Defined by Predicted CATE)

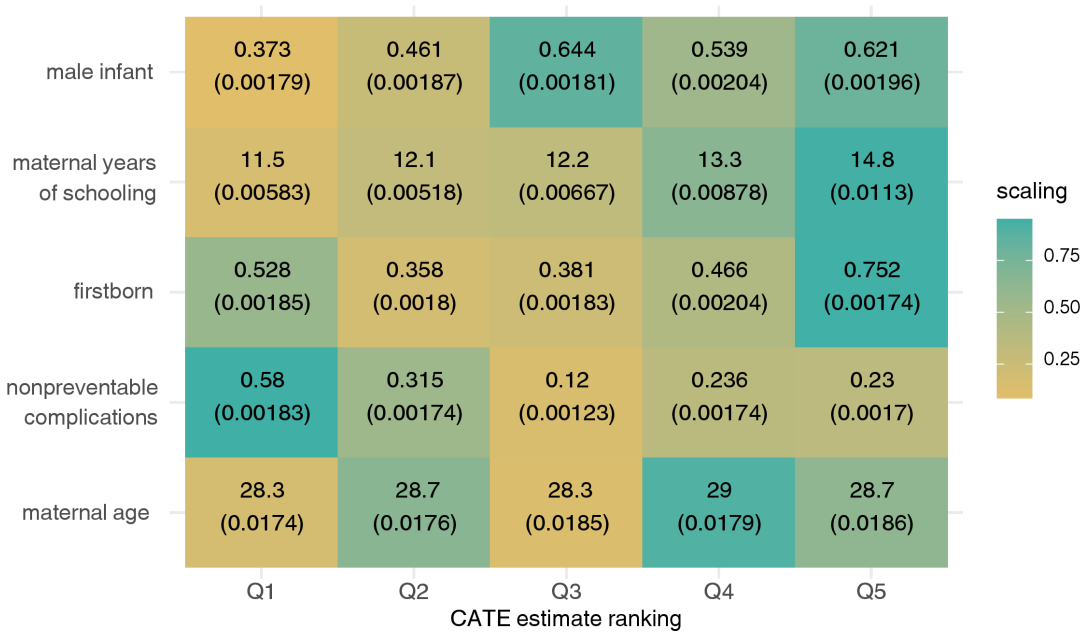


Figure A5.2: Causal Forest Average Covariate Values Within Group Based on CATE Estimate Ranking

Table A6.1: References for ICD-9-CM Diagnosis Code Classification

Diseases	ICD-9-CM Codes
Panel A. Preventable Complications	
Infection	647.0–647.9
Antepartum Hemorrhage	641.8, 641.9
Obstetric Injury	654.7, 664.0–665.9, 669.5
Panel B. Non-preventable Complications	
Malpresentation / Breech Presentation	652.0–652.9, 669.0–669.1
Cephalopelvic Disproportion (CPD)	653.0–653.9
Placenta Previa	641.0, 641.1
Umbilical Cord Complications	663.0–663.3
Obstructed Labor / Dystocia	660.0–662.3
Abnormality of Organs and Soft Tissues of Pelvis	654.0, 654.6–654.64
Previous Cesarean Delivery	654.20–654.23
Panel C. Other Maternal Conditions & Risk Factors	
Hypertensive Disorders of Pregnancy	642.0–642.4, 642.7, 642.9
Severe Pre-eclampsia/Eclampsia	642.5, 642.6
Abnormal Glucose Tolerance (Gestational Diabetes)	648.8

Notes: Codes follow ICD-9-CM (International Classification of Diseases, 9th Revision, Clinical Modification). Ranges denote all valid subcodes within the stated interval (e.g., 652.0–652.9 includes 652.0, 652.1, etc.). Antepartum hemorrhage (Panel A, 641.8–641.9, "other" and "unspecified" causes) is coded separately from hemorrhage attributable to placenta previa or abruptio placentae (Panel B, 641.0–641.2), reflecting the distinction between preventable and non-preventable causes of antepartum hemorrhage described in Section 3.2.

A6 Diagnosis Code Definitions

Table A6.1 reports the ICD-9-CM diagnosis codes used to classify preventable maternal complications, non-preventable maternal complications, and other maternal conditions and risk factors throughout the paper. Panel A lists the codes underlying our composite preventable-complications indicator (infection, antepartum hemorrhage, and obstetric injury), discussed in Section 3.2. Panel B lists the codes for non-preventable complications. Panel C lists the codes for hypertensive disorders of pregnancy and diabetes mellitus in pregnancy, the same definitions underlying the outcomes examined in our mechanism analysis in Table 6 (Section 7).

A7 Assessing the Exclusion Restriction: Bias Decomposition and Additional Tests

The validity of our instrumental variables strategy rests critically on the exclusion restriction: that SARS exposure affects birth and maternal outcomes only through its impact on prenatal care utilization, not through other direct channels. While this assumption is fundamentally untestable, we can assess its plausibility through theoretical reasoning and the empirical validation exercises presented in Section 6. Here we formalize the exclusion restriction, characterize potential violations,

and derive the resulting bias in our 2SLS estimates.

Referring to the reduced-form equation, Equation (5), the exclusion restriction is violated if $SARS$ affects Y through channels other than prenatal care adequacy, $Kessner$. To allow for such violations, suppose the error term in Equation (5) can be decomposed as

$$\upsilon_i = \gamma\, SARS_i + \varepsilon_i, \qquad E[SARS_i\, \varepsilon_i \mid X_i, \kappa, \lambda, \mu, \nu, \tau t_i] = 0, \qquad \gamma \neq 0.$$

Under this formulation, Equation (5) can be rewritten as the "true" outcome equation:

$$Y_i = \beta_0 + \beta_1\, Kessner_i + X_i'\omega + \kappa_{age(i)} + \lambda_{m(i)} + \mu_{p(i)} + \nu_{h(i)} + \tau t_i + \gamma\, SARS_i + \varepsilon_i, \tag{7}$$

where the term $\gamma\, SARS_i$ captures the direct effect of SARS exposure on the outcome and thus represents the exclusion-restriction violation.

With a single instrument $SARS$, the large-sample 2SLS estimand is given by

$$plim(\hat{\beta}_{1,2SLS}) = \beta_1 + \frac{\gamma}{b_1}, \tag{8}$$

where $b_1 (< 0)$ is the coefficient from the first stage regression from Equation (4). The sign and magnitude of this bias depend critically on γ, the direct effect of SARS exposure on outcomes. We consider three potential channels through which SARS could directly affect maternal and infant health, independent of prenatal care utilization.

Maternal Stress Pathways.—The SARS outbreak generated widespread fear and psychological distress that could directly harm maternal and infant health through well-established biological pathways. These mechanisms worsen birth outcomes directly, independent of prenatal care utilization. If stress channels dominate, $\gamma > 0$, implying $\gamma/b_1 < 0$.

Reduced Utilization of Other Healthcare Services.—SARS may have reduced healthcare utilization more broadly, beyond prenatal care specifically, including outpatient visits, specialist consultations, and elective procedures, in ways that independently affected maternal and infant health. If this channel operates similarly to the stress pathway, by removing access to care that would otherwise have improved outcomes, it implies $\gamma > 0$ and $\gamma/b_1 < 0$, in the same direction as the maternal stress channel above.

Precautionary and Compensatory Behaviors.—Pregnant women may have responded to SARS by adopting protective behaviors that directly improve outcomes, selectively avoiding hospitals except for emergencies while simultaneously improving home health practices. If these compensatory behaviors dominate, $\gamma < 0$, implying $\gamma/b_1 > 0$.

In the maternal stress and reduced healthcare utilization channels, our 2SLS estimates would overstate the protective effects of prenatal care; in the precautionary behaviors channel, they would understate them.

We present two additional validation results here, which are discussed in Section 6. Appendix Table A7.1 reports estimates of SARS exposure and prenatal care adequacy on two negative control

Table A7.1: Negative-Control Tests: Infant Sex and Nonpreventable Complications

	(1)	(2)	(3)	(4)
	Male Infant		Nonpreventable Complications	
	Reduced-Form	IV Estimation	Reduced-Form	IV Estimation
Exposure fraction	-0.0037		-0.0152	
	(0.0056)		(0.0145)	
Kessner adequacy		0.0132		0.0538
		(0.0197)		(0.0515)
Risk factors	Yes	Yes	No	No
Infant sex	No	No	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes
Birth month FE	Yes	Yes	Yes	Yes
Mean of dep var	0.524	0.524	0.302	0.302
Observations	335,674	335,674	335,674	335,674

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: In Columns (1) and (2), we exclude the indicator for male infants from the full set of controls. In Columns (3) and (4), we exclude risk factors from the full set of controls.

Table A7.2: IV Estimation Testing Heterogeneous Effects on Women With Hospitalization Records

	(1)	(2)	(3)	(4)	(5)
	Prev Complications	SGA	LBW	Neonatal Mortality	Infant Mortality
Hosp × Kessner	0.0044	-0.0012	-0.0084	-1.3054	-2.9431
	(0.0216)	(0.0221)	(0.0174)	(2.8369)	(4.9795)
Risk factors	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes
Birth month FE	Yes	Yes	Yes	Yes	Yes
Mean of dep var	0.131	0.118	0.049	1.412	3.057
Observations	335,674	335,674	335,674	335,674	335,674

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

outcomes, infant sex and non-preventable maternal complications, both of which are statistically indistinguishable from zero. Appendix Table A7.2 reports the interaction-term test of whether mothers with a prior hospitalization history exhibit differentially larger treatment effects; the interaction coefficient is statistically insignificant across all outcomes.

A8 Preventable Complications by Delivery Mode

This appendix reports the estimates underlying the delivery-mode discussion in Section 7. As shown there, adequate prenatal care raises the probability of cesarean delivery by 7.2 percentage points, and a mechanical benchmark holding within-mode complication rates fixed indicates that this shift in delivery-mode composition can account for roughly 40 percent of the 3.44-percentage-point headline reduction in preventable complications (Table 4, column 2). To assess whether the remaining effect reflects more than delivery-mode routing, Table A8.1 re-estimates the same specification—OLS, reduced-form, and IV, with the full set of risk-factor and fixed-effects controls—separately by realized delivery mode.

The IV estimate for the full sample reproduces the Table 4 headline effect. Restricting the sample to vaginal deliveries, adequate prenatal care reduces preventable complications by 3.86 percentage points, an effect similar in magnitude to, and slightly larger than, the full-sample estimate and statistically significant. Among cesarean deliveries, by contrast, the point estimate is small and statistically indistinguishable from zero, and a Chow test rejects equality of the vaginal- and cesarean-delivery IV estimates (p = 0.02). Because delivery mode is itself affected by prenatal care, the vaginal-delivery estimate should not be interpreted as an effect that can be added to, or subtracted from, the 40 percent mechanical benchmark above. Subject to that caveat, the pattern indicates that the reduction in preventable complications is not confined to the extensive margin of delivery-mode routing: adequate prenatal care also lowers complications among women who ultimately deliver vaginally.

Table A8.1: Effects of Prenatal Care Adequacy on Preventable Complications by Delivery Mode

	(1) All obs	(2) Vaginal	(3) C-section
OLS	0.0095***	0.0088***	-0.0020***
	(0.0010)	(0.0012)	(0.0005)
RF	0.0097**	0.0108**	0.0014
	(0.0030)	(0.0035)	(0.0015)
IV	-0.0344**	-0.0386**	-0.0048
	(0.0107)	(0.0128)	(0.0052)
Mean of dep var	0.1306	0.1957	0.0063
Observations	335,674	220,275	115,395

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: OLS = ordinary least squares estimate; RF = reduced-form estimate using the instrument (exposure fraction) directly; IV = instrumental-variables (2SLS) estimate of Kessner adequacy. All specifications include the full set of risk-factor controls and hospital, zip-code, and birth-month fixed effects, matching column (2) of Table 4. The "All obs" column reproduces the full-sample estimate from Table 4 for reference; "Vaginal" and "C-section" restrict the estimation sample to births realized through each delivery mode.