NBER WORKING PAPER SERIES

EFFECTS OF WIC ON BIRTH OUTCOMES:
EVIDENCE FROM NEWLY DIGITIZED DATA FROM THE NATIONAL ARCHIVES

Marianne Bitler
Danea Horn
Esra Kose
Maria Rosales-Rueda
Arian Seifoddini

Working Paper 34521
http://www.nber.org/papers/w34521

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2025

This research was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD0938898. The findings, conclusions, views, and opinions are solely those of the authors and do not necessarily represent the views of the National Institutes of Health, the Federal Trade Commission or any individual commissioner, or the U.S. government. Thanks for helpful comments to Anna Aizer, Gabriella Conti, Janet Currie, Hilary Hoynes, Margaret McConnell, Katherine Meckel, Jenna Nobles, Zhuan Pei, and Laura Wherry; to seminar participants at UC Davis, UC Irvine, the University of Oregon, the University of Michigan, the University of Delaware, and the Upjohn Institute; and to conference participants at the NBER Children's meeting, ASHEcon, APPAM, HERO, and the VEPPS seminar. Thanks to Hilary Hoynes, Marianne Page, and Ann Stevens for generously sharing data from their paper. We would also like to thank Kathya Tapia-Schythe for outstanding research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Effects of WIC on Birth Outcomes: Evidence from Newly Digitized Data from the National Archives
Marianne Bitler, Danea Horn, Esra Kose, Maria Rosales-Rueda, and Arian Seifoddini
NBER Working Paper No. 34521
November 2025
JEL No. H53, I18, I38

ABSTRACT

This paper examines the effect of nutrition—delivered during pregnancy via the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)—on infant health. We contribute to the literature by adding new data on WIC rollout from the National Archives, employing novel methods, investigating new mechanisms, and analyzing previously unexplored infant outcomes, including the distribution of birth weight. We find robust evidence that access to WIC in a woman's county of residence by the start of their second trimester led to a reduction in the probability of an infant being of low birth weight and to a higher average birth weight, with larger in magnitude effects for infants with Black mothers than for those with White mothers. Our results suggest a lower bound for treatment-on-the-treated effects for low birth weight of -0.9 percentage points and for birth weight of around 33 grams for infants with Black mothers, both economically significant. For White mothers, the effects on low birth weight are much smaller in magnitude and statistically insignificant. These findings show WIC reduced overall disparities between Blacks and Whites in low birth weight by 3.6 percent.

Marianne Bitler
University of California, Davis
Department of Economics
and NBER
bitler@ucdavis.edu

Danea Horn
University of California, San Francisco
danea.horn@ucsf.edu

Esra Kose
University of California, Merced
esrakose@ucmerced.edu

Maria Rosales-Rueda
University of Delaware
and NBER
mrosales@udel.edu

Arian Seifoddini
Federal Trade Commission
seifoddini@ucdavis.edu

1 Introduction

Health disparities among racial and ethnic groups are a pressing concern to researchers and policymakers. These disparities begin at birth and persist throughout the lifetime. This is important because an individual's health at birth is a key determinant of long-term human capital, well-being, and economic self-sufficiency (Currie, 2011; Black et al., 2007). During the 1960s and early 1970s, the United States (U.S.) established a series of means-tested safety-net programs aimed at helping disadvantaged individuals; including Food Stamps, the School Breakfast Program, Head Start, and Medicaid; as well as the program that is the focus of our attention, the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC).

The WIC program began distributing benefits at the tail end of the War on Poverty in January 1974. It currently offers benefits to children aged zero to four and to pregnant, postpartum, or breastfeeding women whose income is below 185% of the federal poverty guideline. During the 1970s, at the time of our study, conditional on living in a location authorized to be served by a local WIC program, eligibility required the presence of a pregnant woman and/or children under 4 (and, later, 5) in the household, that this person be at nutritional risk, and that they participated in or were eligible for a free or sliding-scale medical program. WIC was operated by non-profit or government-operated clinics, and nutritional risk was determined by health professionals while other staff assessed whether families met the means-testing requirements.[1]

WIC has the potential to influence health at birth and beyond, to improve nutrition-related health outcomes, and reduce outcome gaps between groups. First, it provides large amounts of nutritious food rich in specific nutrients frequently missing from the diets of low-income families at the time. WIC was used by a large share of pregnant women; in 1986 about 36% of Black pregnant women and 15% of White pregnant women paticipated in the program. At the start of WIC, in most locations, participants received vouchers that could be redeemed at participating stores for specific food items. In limited instances, participants had the food items delivered to their homes, or they were able to pick up those items at local warehouses or, rarely, at the clinics themselves. The research literature on WIC has established that the current version of this program improves the quality of nutritional intake and birth outcomes (e.g., Figlio et al., 2009; Rossin-Slater, 2013; Currie and Rajani, 2015; Currie and Rossin-Slater, 2020).

[1]Later on, more explicit rules about family income being below a particular threshold were put in place, matching the limit used for the reduced-price school meals programs.

We add to this rich WIC literature by examining the rollout of the WIC program during the 1970s in combination with Vital Statistics Natality records. Shortly after WIC's introduction, USDA enacted a series of program changes, but was also constrained by lawsuits brought by advocates and was forced to expand the program significantly. We combine this quasi-experimental variation in WIC's implementation timing with Vital Statistics birth records to study its effects on pregnancy and birth outcomes. Our paper builds on the empirical foundation laid by Hoynes et al. (2011),[2] who examined the effect of WIC adoption on health at birth using two-way fixed effects (TWFE) models. We have expanded their rollout data with extensive new information on county's WIC adoption from National Archives records of State WIC plans through 1980 which we append to HPS's rollout data, which ended in 1979. These new data allow us to detect with more granularity the timing and geographic coverage of WIC's introduction and to uncover nearly 450 additional counties that had WIC present.

Our empirical strategy leverages staggered temporal and geographic variation in the introduction of WIC, using the new difference-in-differences methods by Sun and Abraham (2021), which are robust to the presence of treatment effect heterogeneity across places and over time, as well as to possible contamination of some estimates by the effects from other periods. Our preferred estimates use inverse propensity score weights to adjust the control group so that pre-WIC across-county averages of social, demographic, and economic conditions are balanced across the treatment and control counties. We additionally combine ITT estimates with an estimate of the effects of rollout on participation to produce bounds on the average treatment effect on the treated. We pay particular attention to how the program influences birth outcomes by race, focusing primarily on Black and White infants.[3] We focus on the heterogeneous impact of WIC by race because racial disparities in infant health have been persistent over time (e.g., Vu et al., 2024; Kennedy-Moulton et al., 2022; Owens-Young and Bell, 2020; Rauscher and Rangel, 2020; Singh and Yu, 2019; Chambers et al., 2018). In addition, we explore mechanisms that could explain this heterogeneity.[4]

The results indicate that the adoption of the WIC program is associated with a 0.31 per-

[2] We use "HPS" to reference this important work.

[3] We exclude Native Americans from this study because Native Americans who resided in Indian Land Areas often had access to WIC run by Indian Tribal Organizations. These entities may have operated differently from the clinics and local agencies that operated the program elsewhere. Further, much of the rest of the social safety net in these lands is distinct from that offered in the rest of the U.S..

[4] Malinovskaya (2023) looks at long-run effects of WIC exposure in utero on adults, using a combination of the HPS WIC rollout and more detailed information on Texas combined with restricted-use Census data. She finds significant, economically meaningful effects for Black adults' health (concentrated in places with higher teen pregnancy rates and poverty rates), and similarly large effects for White adults' self-sufficiency in Texas.

centage point reduction in the incidence of low birth weight among infants with Black mothers (henceforth "Black infants"), along with a 12-gram increase in average birth weight, a 0.38% increase over control group mean birth weight of 3080 grams; both estimates are statistically significant and economically meaningful. This compares to a statistically insignificant decline in the incidence of low birth weight of 0.06 percentage points and a statistically significant increase in average birth weight of 5 grams, a 0.14% increase over control group mean birth weight of 3340 grams, for infants with White mothers (henceforth, "White infants"). These results suggest that WIC led to a decline of about 0.25 percentage points ($\approx$ 3.6 percent) in the Black-White low birth weight gap, and a decline of around 6.6 grams on average in the Black-White birth weight gap ($\approx$ 2.5 percent).[5] We further investigate the impacts of WIC at different points in the birth weight distribution and find that the effects are large and economically meaningful at the lower end, mainly for Black infants. WIC shifted the entire Black infant birth weight distribution upward.

We consider multiple plausible explanations of why WIC improves infant health outcomes for infants of Black mothers more than for infants of White mothers. First, eligibility, participation, and, to a lesser extent, take-up of WIC were substantially higher for Blacks than Whites. Second, Black infants in this period had far worse baseline (pre-WIC rollout) birth outcomes than their White counterparts.[6] Third, infants of Black and White mothers had varying levels of individual risk factors. For example, Appendix Table A.1 reports that the teen Black birth rate was about double the teen White birth rate during the study period, and it is well-known that teen births are riskier than other births. Similarly, single motherhood was more common for Black mothers than White mothers, with about 6 times as many Black births being to single mothers in this time. Fourth, differences in place-based factors faced by infants of Black mothers versus those of White mothers could also play a role. For example, patterns of residential segregation suggest that places where Black individuals lived might differ from those where White individuals lived, and this could affect birth outcomes differentially. Overall, we find that individual and place-based factors play little role in explaining the differences in the effects of WIC by race. The race differences in WIC impacts seem to be explained by (1) racial differences in WIC program take-up and eligibility, and (2) baseline differences in outcomes by race, leading to higher returns on this investment for Black mothers relative to White mothers.

[5]To calculate this, we divide the differences of the estimated effects for Black and White infants by the corresponding differences in their control means. We calculate the reported effect for low birth weight: $100 * \frac{(-0.3078-(-0.0585))}{(13.434-6.432)} \approx -3.6$, and for birth weight: $100 * \frac{(11.7089-5.1234)}{(3080.080-3339.085)} \approx -2.5$.

[6]Appendix Figure A.1 shows that infants of Black mothers had much lower average birth weights in 1970 (before the introduction of WIC) than did those of Whites.

Our study contributes to several strands of literature. First, it adds to the growing body of research on War on Poverty programs, which collectively sought to improve the well-being and economic prospects of low-income families through investments in health, nutrition, and early childhood development. A rich set of studies has examined the long-run and contemporaneous impacts of initiatives such as Head Start, Community Health Centers, Food Stamps, and Medicaid on child and adult outcomes (e.g., Ludwig and Miller, 2007; Hoynes et al., 2016; East et al., 2023; Kose et al., 2024). We focus on WIC, one of the last major safety-net programs created in this era that targeted nutritional investments specifically to pregnant women and young children, a group often at high risk for nutritional deficiencies. By doing so, we provide new evidence on how this nutrition intervention complemented the broader War on Poverty policy landscape.

Finally, our paper extends the literature on the long-term consequences of early-life conditions by examining the role of targeted nutrition support during pregnancy in reducing racial disparities in birth outcomes. Prior work has established that improvements in the prenatal environment can have lasting effects on health, education, and labor-market outcomes (e.g., Currie, 2011; Almond et al., 2018). We contribute to this literature by showing that WIC's introduction during the 1970s led to measurable improvements in infant health and by documenting that these benefits were disproportionately larger for Black infants, narrowing the Black–White birth outcome gap.

The rest of the paper is organized as follows. In section 2, we discuss background information about the WIC program and its rollout, and section 3 describes our data sources. Section 4 presents our empirical strategy and discusses the underlying identification assumptions. In Section 5.1, we present our results, and Section 7 concludes.

2 The WIC program and WIC rollout

Background

The WIC program, at its inception in 1974, gave recipients either vouchers to purchase healthful foods rich in nutrients of concern or access to home-delivered food or to food at warehouses or clinics. These benefits were available to categorically eligible pregnant, postpartum, and breastfeeding women and young children. Eligibility was conditioned on using or being eligible for free or subsidized health care, while also being nutritionally at risk. Participants also had to live in the "area served" by the relevant local WIC program.

Importantly, WIC benefits accounted for a significant share of essential foods and nutrients for pregnant women. The WIC "food package" for pregnant women in 1976 consisted of foods high in important nutrients, including milk or cheese, eggs, high-iron breakfast cereal, and high-vitamin fruit juice (Bendick Jr. et al., 1976). In 1976, in California, WIC benefits for pregnant women included 60–100% of the recommended dietary allowance (RDA) for protein, calcium, iron, vitamin C, and vitamin A, contributing around 25% of the calories recommended for pregnant women.[7] These nutrients have been identified as critical for the healthy development of the fetus during pregnancy (Shapira, 2008; Marshall et al., 2022). WIC also provided a substantial amount of food, with average monthly food benefits of about $20 per month, or about 39 percent of average food at home spending in 1974 (Manchester, 1976).

rollout of WIC between 1974 and 1980

The WIC program was established by law in 1972; however, the first clinic only became operational in January 1974 in Pineville, Kentucky. A series of lawsuits from outside groups and efforts by the Nixon and Ford administrations to shrink, modify, or terminate the WIC program led to considerable variation in when and where WIC programs were established. The WIC program has never been an entitlement, and the process for selecting sites varied extensively across the early period of its rollout. Numerous congressional hearings include presentations from non-profits that wanted to establish WIC clinics but were not selected. This is important for our key identification assumption that the times and locations of WIC rollout were idiosyncratic, conditional on the pre-WIC characteristics of counties. Below, we explore the extent to which the timing and location of WIC clinic establishment are explained by counties' economic, demographic, or infant health conditions or by the presence of the other War on Poverty programs before the program was introduced.

We build upon WIC rollout data originally collected by Hoynes et al. (2011), who assembled information on the year of the first WIC program implementation in a county from four sources: i) A list of local agencies from 1974 Congressional Hearings; ii) a list of local agencies from a 1975 WIC report by the Children's Foundation; iii) a list of counties served by WIC programs in 1978 in 13 states that participated in the Migrant WIC program (WIC Migrant Directory); and iv) a list of local agencies operating in 1979 from the Children's Foundation directory of WIC locations.

[7] The recommended dietary allowance (RDA) is the "Average daily level of intake sufficient to meet the nutrient requirements of nearly all ... healthy individuals; often used to plan nutritionally adequate diets for individuals." See https://ods.od.nih.gov/HealthInformation/nutrientrecommendations.aspx for more information.

We fill in many gaps in the data used by Hoynes et al. (2011) (HPS) by using National Archives sources and some other documents.[8] We digitized information from WIC Program State plans that states had to provide to USDA from fiscal year 1978[9] to 1981.[10] In 1976, a court order (based on a lawsuit that USDA lost) required the agency to ensure that the program was extended to those women and infants most in need. As a result, USDA required State agencies to develop Affirmative Action Plans as part of their WIC program's State Plans, which provided us with the records of WIC agencies and, in most cases, their serving counties in a given fiscal year. We use these sources of information to construct WIC implementation data.

Each state published lists and/or maps that listed WIC agencies currently participating in a given fiscal year and counties served (the agency's county and any other counties to which the agency provided WIC benefits). They also listed agencies approved and planning to participate and their expected counties served, as well as counties not participating in the program. Some states included information by county concerning special populations served (migrants and Native Americans).

The information included in these reports and how it was presented varied widely by state. We digitized these State Plans and some other documents to measure WIC rollout. We combined our newly digitized data with the digitization of the HPS sources to create a measure of WIC rollout that corresponds to the month and year of the first WIC program implementation in a county. In addition, for January 1980, we have data on take-up from administrative caseload counts and predicted eligibility counts at the county level from a 1981 Congressional hearing. Data Appendix A describes our data and the digitization process in detail.

Panel (a) of Figure 1 shows the evolution of WIC adoption over time among the states we study.[11] The figure shows that in 1974, WIC started in 356 counties, and its adoption

[8] HPS have limited data for 1976 and 1977, but we also have less granularity in those years. This is due in part to the fact that 37 states and DC were not participating in the Migrant WIC Program and, hence, were not in the WIC Migrant Directory and have no information in HPS's data for those years. It is also attributable to the fact that WIC state plans were not available prior to FY 1978. States were required to create these plans because USDA lost a lawsuit and was mandated to both open the program more broadly to communities with needy applicants and to document county need in the State Plans.

[9] We have State Plans for seven states in FY 1978.

[10] The Federal Government's fiscal year currently runs from October 1st to September 30th, starting Fiscal Year 1977. During our sample period, the Fiscal Year moved by one quarter, so the Fiscal Year ended earlier in the year, on June 30 through Fiscal Year 1976.

[11] We exclude Virginia, Alaska, Hawaii, and other U.S. territories which have WIC (American Samoa, Guam, the Northern Marianas, Puerto Rico, and the Virgin Islands) from our sample due to missing county characteristics (HI and other territories) and changing geography (VA, AK). We also note the territories have a very different public safety net than do U.S. states. We do not have comprehensive rollout documents

expanded significantly throughout the rest of the 1970s, with 2,376 counties (out of 2992) having adopted by 1980. Panel (b) of Figure 1 presents the geographic variation in WIC adoption timing, which shows that there is considerable variation across counties and over time within and across states. We capitalize on this variation to examine the effect of WIC on infant health outcomes.[12]

3 Data

Data on birth outcomes come from the Vital Statistics birth records from 1968 to 1981.[13] These data are compiled from individual birth certificate records and represent either a 50% or 100% sample of births occurring in each state. Our primary outcomes of interest are whether the infant is considered low birth weight (birth weight $< 2,500$ grams, an important threshold at the time) and birth weight. We also examine several additional outcomes from the birth certificate: Whether the infant was classified as very low birth weight (birth weight $< 1,500$ grams), very high birth weight (birth weight $> 4,500$ grams), and whether the infant was small or large for gestational age;[14] as well as whether the infant was born prematurely (before 37 weeks of gestation). The prenatal care measure we use is constructed using the Kessner index for adequacy of prenatal care.[15] Lastly, we explore the birth weight distribution in more detail, considering the probability that birth weight is under a range of cutoffs from 1500 to 4500 grams.

The birth certificate data also include information on mother's age and race and place

from 1976 and 1977, and the number of counties in those years is based on retrospective information in some of the State WIC plans. However, we do not think this is important as other documents show that rollout in those years was relatively rare (National Research Council and Division of Behavioral and Social Sciences and Committee on Child Development Research and Public Policy and Panel for the Study of the Policy Formation Process, 1982).

[12]Appendix Figure A.2 compares our compilation of WIC rollout with that from HPS, showing counties with WIC locations (around 450 additional counties) relative to this preexisting literature due to our novel sources for rollout. In addition, Appendix Table A.2 reports cross-tabulations of adoption timing by counties across two data sources. In the diagonal entries, both sources report the counts of counties that adopted WIC in the same year and the off-diagonal entries show the difference in the timing of WIC adoption in our data compilation versus HPS's data.

[13]Vital Statistics micro-data are available starting in 1968, and county identifiers are publicly available during our study period.

[14]"Small for gestational age" is defined as being in the bottom 10% in birth weight of all infants born at the same gestational age in the same year, while "large for gestational age" is defined as being in the top 90% among infants born at the same gestational age.

[15]The Kessner index for adequate prenatal care is based on the gestational age and the number of prenatal visits. For babies under 14 weeks of gestation, at least one prenatal visit is considered adequate. Between 14–17 weeks, at least two visits; 18–21 weeks, at least three visits; 22–25 weeks, at least 4 visits; 26–29 weeks, at least 5 visits; 30-31 weeks, at least 6 visits; 32-33 weeks, at least 7; 34-35 weeks, at least 8; and 36 weeks or more, at least 9 visits are considered as adequate.

of residence; child gender; and, in some states and years; marital status, and mother's educational attainment. More details about our sample selection are available in Data Appendix B. The data are collapsed into cells by county by month-and-year of the start of the second trimester of birth (the most common time for women to join WIC), and by race of the mother. Our results focus primarily on infants of Black and White mothers, but we do present some results for an "Other race" group in the Appendix.

Regressions and summary statistics are then either weighted by the number of births in each cell (race by county by month and year of the second trimester), or, in our preferred specifications, by the number of births times an adjustment for the control group to balance important pre-WIC county characteristics across the treatment and control groups as discussed in Section 4 in more detail.[16] We do this cell-based analysis to facilitate computation and, more importantly, because this is the relevant level of aggregation for the difference-in-differences (DD) models we use, which allow for treatment effect heterogeneity across counties or over time by race group.[17]

We combine these Vital Statistics data with the WIC rollout information described above and various other county-level characteristics, including county population in 1970; exposure to other important food assistance and War on Poverty programs shown to affect birth outcomes such as Food Stamps, Community Health Centers (CHC), Commodity Distribution Program (CDP), and Commodity Supplemental Food Program (CSFP); and economic, safety net, and demographic characteristics such as spending on the Medicare program and unemployment insurance (a proxy for the weakness of the county-level labor market).

We assign exposure to WIC based on the calendar month and year of the start of the second trimester of each birth, using the reported gestational age where available and imputing a value of 40 weeks when gestational age is missing. We use this relatively conservative timing to assign WIC adoption because in our time frame, women typically become aware of their pregnancy status after five weeks of gestation and thus are unlikely to enroll in the WIC program until then. To calculate the month and year of the second trimester, we first compute the month and year of conception using information on the month and year of birth and gestational age in weeks (converted to months). Approximately 30% of births are missing gestational age. For these individuals, we use the population median value of 40

[16]We construct the measure of the number of births in a cell by summing up the total number of records in the county, time, and race cell (if it is a 100% reporting county) or by 2 times the number of total records in the county, time, and race cell (if it is a 50% reporting county).

[17]Our results are also robust to using traditional TWFE models, individual-level data, and diverse sets of controls and specifications.

weeks to impute the missing value.[18] After calculating the month and year of conception, we calculate the month and year of the start of the second trimester. Our results are also robust to assigning 38 or 39 weeks of gestation for those births with missing data and using the year of the third trimester to assign exposure to WIC.[19] For time-varying covariates measured at the year level, we match them according to the year of the second trimester.

Our analysis data include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We present summary statistics for outcomes and control variables in Appendix Table A.3. On average, Columns 1 and 2 of this table show that Black infants are more likely to experience adverse birth outcomes, such as low birth weight, being small for gestational age, preterm birth, and overall lower birth weight compared to White infants. Additionally, Black mothers are more likely to reside in counties with earlier access to Food Stamps and CHC programs, leading to greater exposure to these initiatives than White mothers. Furthermore, Black mothers are disproportionately located in counties with higher concentrations of poverty. These patterns persist in both WIC and non-WIC areas (Columns 3-6 of the table). These findings underscore the importance of analyzing the effects of WIC by race, suggesting that nutritional interventions could disproportionately benefit Black infants, who face a larger baseline disadvantage, in line with the human capital theory.

4 Empirical strategies

We exploit temporal and geographic variation in the rollout of WIC to estimate the causal effect of access to nutritional investments during pregnancy on birth outcomes, using a difference-in-differences (DD) strategy. To address potential bias in staggered DD designs when there is heterogeneity in treatment effects across locations over time, we use the Sun and Abraham (2021) approach, utilizing "never-treated counties" during our time span as the control group.[20] As explained in detail below, to address the issue that our never-treated counties may be different in observable ways from those that adopted WIC during our study period, we use a propensity score reweighting approach.

The standard DD specification is as follows:

[18] The bulk of this missing gestation is because not all states reported detailed gestational age on the birth certificate in all the years in our data, although a number of birth certificate records are simply missing gestational age.

[19] We discuss these results in Section 5.4 and present these results in Appendix Table A.11.

[20] In our setting, the "never treated counties" primarily adopted WIC after 1980.

$$Y_{rct} = \alpha + \beta \cdot WIC_{ct} + \gamma_c + \tau_t + \epsilon_{rct}, \tag{1}$$

where Y_{rct} is the birth outcome by cell defined at the r race, c county, and t month and year of the start of the second-trimester level. WIC_{ct} is an indicator equal to 1 if individuals are exposed to WIC based on their county of residence during the month and year of the start of their second trimester.[21] γ_c and τ_t are county and month-by-year of the start of the second trimester fixed effects, which absorb time-invariant county characteristics that could be correlated with WIC rollout and time-varying national changes, respectively. In our preferred specification, each cell observation is weighted by the number of births in the cell for the treated group and by the number of births times the propensity score adjustment for the control group. Our estimates are generally stratified by race, with a focus on Black and White mothers.[22] Standard errors are clustered at the county level. The key coefficient of interest is β, which measures the static DD estimate of the effect of WIC. This coefficient is identified from changes in birth outcomes between births exposed to WIC and those not exposed before and after the rollout. Lastly, to examine dynamic treatment effects, we estimate an event-study version of Equation 1.

4.1 Determinants of WIC adoption

The validity of our research design relies on the assumption that the timing of WIC rollout across counties is uncorrelated with other county, time-varying determinants of our outcomes of interest. To provide evidence for this assumption, we test whether 1970 or earlier county characteristics (pre-program) predict the timing of a county's WIC program implementation following a similar approach to that used by other War on Poverty studies (e.g., Hoynes and Schanzenbach, 2009; Bailey, 2012; Bailey et al., 2022). The dependent variable is time to adoption after 1973 (the calendar year of WIC's introduction minus 1973). The results presented in Appendix Table A.4 include models with (even Columns) and without (odd Columns) state-fixed effects, with results that are both unweighted (Columns 1–2 and 5–6) and weighted by the number of births in each county in 1970 (Columns 3–4 and 7–8). We find that county characteristics like urbanicity/rurality, population, the poverty level, the number of hospitals per capita, the number of active medical doctors per capita, and the presence of a CHC before 1974 are all associated with earlier WIC introduction. Columns 5–8 also consider the importance of two commodity programs in use during this period: the Commodity Distribution Program (CDP), which was a broad based food program providing

[21] Our results are robust to defining exposure according to the start of the third trimester of pregnancy.

[22] We also present the results for non-Native "Other" mothers.

foods to recipients that typically ended when the Food Stamp program became available in a county; and the Commodity Supplemental Food Program (CSFP), a targeted food program offering foods to a similar population to that participating in WIC.[23] Note here that the quantitative importance of these predictors is small, and most of the variation remains unexplained, as illustrated by the relatively low R-squared, which ranges between 0.1 and 0.4.

USDA required states to create an Affirmative Action plan, which they had to report in their State Plans and were supposed to use in deciding where to expand WIC after the lawsuit they lost in 1976. Despite this, there is no evidence that either the share of non-White persons in the county or the infant mortality rate predicts WIC adoption timing in our sample period. The coefficients are small and statistically insignificant. This supports our choice of identification strategy, as there is no substantive evidence of selection on these observables in the timing of WIC rollout. However, there still may be concerns about selection into timing beyond the time invariant local factors removed via the inclusion of county fixed effects and the national shocks removed by the year effects. To address remaining concerns about why WIC was adopted at particular times in particular locations, we take two approaches: 1) propensity score reweighting as our preferred approach, and 2) flexibly controlling for county characteristics as an alternative.

4.2 Propensity-score weights

First, a credible DD approach relies on the assumption that our adopting counties are comparable to our non-adopting counties, which may be challenging in our setting, as WIC rollout is not completely random. Appendix Table A.5, Column 1 shows a number of statistically significant imbalances in raw pre-treatment mean characteristics between WIC and non-WIC counties by 1980. These even survive multiple hypothesis testing (see Column 3). Thus, in addition to TWFE and Sun and Abraham (2021) estimates using all counties that adopted by 1980 and all other non-adopting counties (excluding Alaska, Hawaii, and Virginia), our preferred specification uses an alternative approach related to but simpler than Sant'Anna and Zhao (2020)'s doubly robust approach and that used by Cunningham and Goodman-Bacon (2024). We use propensity scores from models that predict whether a county had adopted WIC by 1980 to adjust our cell weights. This approach helps to construct a valid comparison group of never-treated counties by weighting the control counties to have the same observable characteristics on average as the treatment group counties.

[23] At the beginning of WIC, USDA did not allow WIC and the Commodity Supplemental Food Program to operate at the same time, leading to the later adoption in CSFP locations.

We adjust the weights of possible control county cells. This approach involves estimating a logistic model of having adopted WIC by 1980 as a function of pre-1974 county characteristics, which include birth outcomes, the presence of the CHC and Food Stamps programs by 1974, county demographics, measures of the reach of the health care system, and 1970 Census poverty and population. More precisely, the model includes 1970 or 1960 measures of characteristics used to designate counties as medically underserved in the 1980s. These include the share of the population under poverty in 1969, the share of the population 65 and older in 1970, the infant mortality rate in 1970, and the number of doctors per capita in 1960. Moreover, we include additional demographic controls such as the share of the county population under 5 years old, the share of the adult population with various levels of education, the share non-White, the share urban/rural, real spending on Medicare per capita, real spending on State Unemployment Insurance payments per capita, and the natural log of the county population, all measured in 1970. It also includes the teen birth rate in 1970, the share of births with low birth weight in 1970, the presence of commodity programs (CDP and CSFP) and state-level measures of White and Black AFDC participation by women and children.[24] Appendix Figure A.5 presents the distribution of the predicted probability of being treated for both the ever-WIC by 1980 (treated counties) and the never-treated counties, and shows a substantial overlap.

After estimating the predicted probability of being treated, referenced as $\hat{p}$, we compute cell weights as follows. For treated counties, the cell weight is left as is, which is the number of births in the cell. For control counties, we weight the cells by the number of cell births times $\hat{p}/(1-\hat{p})$, where $\hat{p}$ is the county probability of having adopted WIC by 1980.[25] This is our preferred specification and Column 4 of Appendix Table A.5 presents balance tests of pre-program characteristics after re-weighting the comparison counties to match the treated counties. Most of the imbalances from the raw data disappear, and the differences are no longer statistically significant after adjusting for multiple hypothesis testing. Similarly, Column 7 includes state fixed effects (in the balance test) and shows that within states, there are no imbalances (except for the percentage of first-borns in 1970) between treated and never-treated counties.

To review, our preferred specification uses Sun and Abraham (2021)'s approach with propensity score reweighting to account for pre-WIC differences in county characteristics.

[24] The state-level measures of child and women participating in AFDC on the eve of Medicaid rollout were used by Goodman-Bacon (2021) as a measure of the eligible population for Medicaid at its start, as AFDC recipients were all eligible for Medicaid.

[25] The controls in this regression, which include the main controls about the safety net, do not vary over time and are absorbed by the county FE; thus, we do not control for them as is normal in a doubly robust setting. However, we do not do this exercise separately by timing group.

We also test the robustness of our results without propensity score reweighting, both with and without time-varying county-level controls; including CHC exposure, population, and the number of hospitals per capita each year, as well as the Food Stamp Program rollout. We present these results in Section 5.4, paragraph one. Importantly, our conclusions do not rely on the inclusion of these covariates. These estimates are in alignment with our preferred specification and those from other specifications that only include county and year-fixed effects, as demonstrated through a multitude of robustness checks. Also, our findings are robust to further alternative specifications that include 1970 county characteristics interacted with year dummies (e.g., Almond et al., 2018), and even to including state-by-year fixed effects, which control for any other state-wide programs being introduced.

4.3 Parallel trends with p-score weighting

The second key assumption in DD models is the parallel trends assumption. We assume that in the absence of the program, counties with and without WIC would have had similar trends in birth outcomes over time.

Figures 3 and 4 show our dynamic estimation and visual presentation of parallel trends, using the propensity score weighting and Sun and Abraham (2021), for Black and White mothers, respectively. The dots represent the point estimates from the event studies and the bars represent confidence intervals. Overall, we do not find much evidence of pre-trends for Black mothers (except for the Kessner Index of adequate prenatal care) as the estimated coefficients in the periods before WIC are all small and statistically insignificant. However, for White mothers, we find some evidence of differential pre-trends in several birth outcomes, so we interpret their estimated effects with greater caution. In addition, we perform the pre-trends analyses developed by Roth (2022) and consider how coefficients change when we assume various levels of pre-trend deviations as suggested by Rambachan and Roth (2023). We discuss these analyses in more detail in Section 5.1.

4.4 Evidence about negative weights and correlation of weights with timing and other factors

Next, we acknowledge that the standard practice until recently has been to estimate DD models using two-way fixed effects (TWFE) specifications.[26] However, as mentioned above, recent developments in econometrics have shown that this method may lead to coefficients for the ATT to be biased in staggered designs such as ours (Goodman-Bacon, 2021; Sun

[26] We present TWFE estimation results as an alternative to our preferred specification in Section 5.4.

and Abraham, 2021; Callaway and Sant'Anna, 2020; de Chaisemartin and D'Haultfœuille, 2020).

To assess whether treatment effect heterogeneity is a potential issue in our setting, we perform a diagnostic test proposed by de Chaisemartin and D'Haultfœuille (2020). We present these test results in Appendix Table A.7.[27] The TWFE weights are correlated with treatment year, mother's race, hospitals per capita, and state unemployment insurance compensation, indicating the possible presence of treatment effect heterogeneity. Also, around 13% of the TWFE weights are negative. This evidence suggests that TWFE may not be appropriate in our setting because heterogeneous treatment effects can lead to bias.

It is for this reason that we use the procedure developed by Sun and Abraham (2021) (SA) that accounts for dynamic treatment effects across counties that adopted WIC in a given year (cohorts). This methodology divides units into different cohorts depending on when they receive the policy, plus a never-treated group that is used as a comparison group. The estimation procedure identifies the cohort-specific average treatment effect on the treated, $\text{CATT}_{e,l}$, in relative time period l for the cohort that received the intervention at time e. Their interaction-weighted estimator uses a saturated regression, interacting relative time dummies with cohort dummies (excluding indicators for the comparison group). The aggregated event-study coefficients (for each relative time period, l) are computed by taking the weighted average of the $\text{CATT}_{e,l}$ using weights corresponding to the sample shares of each cohort in each relative time period.[28] Sun and Abraham method gives a consistent estimate of the treatment effect under the parallel trends and no anticipation assumptions. For comparison, we present estimations using TWFE in the robustness section. We remind the reader that in our setting, the "never treated group" is the set of counties that had not yet adopted WIC by 1980.

5 Results

5.1 Infant health outcomes

Table 1 presents our static DD effects of WIC on various birth outcomes using Sun and Abraham and the propensity score reweighting approach. These outcomes include the

[27] We note that since we weight our regressions by number of births and adoption increases over time, there is also a potential mechanical correlation with treatment year.

[28] The computation for the static DD estimator is similar, but in that case, there are only two relative-time bins: one with all the pre-periods and the other with all the post-periods. The first is omitted due to collinearity, similar to the omission of -1 in the dynamic estimation, and serves as the comparison point for the differences in average outcomes.

birth being low birth weight (<2500 grams), being very low birth weight (<1500 grams), being small for gestational age (below the tenth percentile of birth weight for gestational age), being preterm (<37 weeks gestation), birth weight (in grams), being high birth weight (>4500 grams), being large for gestational age (above the ninetieth percentile of birth weight for gestational age), and the Kessner index for adequate prenatal care (constructed based on gestational age and the number of visits),[29] in Columns 1 to 8, respectively.[30]

These results show that the intent-to-treat (ITT) effects of WIC's introduction in a county are significant and more meaningful for Black births relative to White births. We find that the improvements in health at birth, measured by the probability of an infant being of low birth weight, being of very low birth weight, or being small for gestational age, are larger in magnitude among births to Black mothers than among births to White mothers (Columns 1–3). Black infants exposed to WIC in their county experienced a statistically significant decline in low birth weight of 0.31 percentage points (Column 1), compared to a decline of 0.06 percentage points for White infants. Relative to the pre-period mean, the results imply ITT effects of -2.3% and -0.9%, respectively.

Turning to the incidence of very low birth weight (Column 2) and small for gestational age (Column 3), we find that relative to the mean, WIC led to a decrease in these outcomes by 5.6% and 4.5%, respectively, for Black babies compared to 1.0% and 1.1% for White babies. Overall, the effects of WIC on improving these negative markers at birth are more than four times larger for Black births than for White births. The medical literature has documented that birth weight and intrauterine growth restriction (commonly proxied by small for gestational age) are responsive to nutritional investments, the focus of the WIC program (e.g., Kramer, 1987a,b).

Regarding the likelihood of preterm birth, we find a positive but insignificant effect of WIC exposure larger in magnitude for Blacks than for Whites (Column 4). In Column 5, our findings indicate that WIC implementation led to a statistically significant increase in birth weight of 12 grams for Black babies compared to an increase of 5 grams for White babies, a 0.38% and 0.14% increase over the control group birth weight, respectively. Also, we find that WIC exposure increases the probability of being of high birth weight (Column 6) and large for gestational age (Column 7), particularly for Black infants. However, these last two outcomes are quite rare, and their importance was less well understood in this time period.

[29]See Footnote 15 for details on how we construct the Kessner index.

[30]In addition, we examine neonatal and infant mortality and find no significant evidence of WIC impacts on those outcomes. The details about this analysis are described in Appendix C.

Overall, our findings suggest that as a nutritional intervention, WIC not only affected birth weight on the left tail (low birth weight) but also moved a small number of births from being in the normal birth weight range to being high birth weight. We explore this further in Section 5.2 below. Lastly, in Column 8, we find a marginally statistically significant positive effect on the Kessner index for adequate prenatal care for Black births.

One might be concerned about multiple hypothesis testing (MHT) since we examine several birth outcomes. Appendix Table A.16 presents the results of such corrections. Column 1 repeats the *p*-values from Table 1. Column 2 presents results using the approach of Holm (1979) that controls the family-wise error rate (the probability that you reject any null when it is true). This shows that for Blacks, most of our significant effects (4 out of 5 outcomes) remained statistically significant after implementing this quite conservative multiple hypothesis adjustment at the 5% level. Column 3 shows the *q*-values that control the false discovery rate (the share of your significant results that is false). This indicates that all of our results for Blacks are robust to this less conservative MHT adjustment.

Next, we use event study specifications to examine the dynamic effects and test the parallel pre-trends assumption. Our results for Black mothers in Figure 3 show (1) evidence suggestive of dynamic effects for multiple outcomes up to 3 years in the post period and (2) no strong evidence of pre-trends in most outcomes. Notably, the figures show a gradual decline in the likelihood that infants are of low birth weight, very low birth weight, and small for gestational age, post-WIC introduction. For birth weight, we observe a gradual increase after the program rollout. For White mothers in Figure 4, the effects post-WIC are smaller in magnitude and noisy, and there is also more evidence of pre-trends.[31]

We further assess the parallel trends assumption and potential violations in Appendix A. First, following Roth (2022), we estimate the slope of the largest linear parallel trend violation in the pre-period that we could detect with various levels of power (50 and 80 percent) and present this in Appendix Figures A.6 and A.7 for a number of outcomes (low birth weight, small for gestational age, very low birth weight, and birth weight), assuming the pre-trends are the least favorable for our hypotheses.[32]

[31] We additionally use this event study specification to examine infant mortality as an outcome. The results are presented in Appendix C.

[32] These potential pre-trend violations are depicted as red (solid) lines in our Appendix event study figures A.6 and A.7. The dashed blue lines illustrate the expected event study coefficients that we would see on average, conditional on not having detected a significant pre-trend when in fact the hypothesized red line was the true pre-trend for these levels of power. This analysis suggests that for Blacks, most of the post-WIC estimates by year after WIC adoption for the outcome small for gestational age (including the confidence intervals) cannot be explained by the expected pattern resulting from this parallel trend violation, as the blue line is outside the CI for the estimated point estimates for periods 2 and 4 plus after WIC adoption for power 0.5. This relatively stringent test is less confirming for low birth weight, but the general pattern

Second, we test for the joint significance of all the pre-treatment coefficients using F-tests. The results are presented in Appendix Table A.6, which includes the usual p-values along with two adjustments for MHT. For Black mothers, we find some evidence of differential pre-trends in the adequate prenatal care indicator, as the joint-significance p-value remains significant at 10% after correcting for MHT. For Whites, the p-values suggest potential pre-existing trends in low birth weight, very low birth weight, overall birth weight, and the adequate prenatal care index.

5.1.1 Treatment-on-the-treated (TOT)

To translate our ITT estimates into back-of-the-envelope average TOT effects, we would ideally have WIC caseload data by county and race, and we would estimate a first stage. No such granular caseload data are available as far as we know, however. Instead, we follow much of the War on Poverty literature (e.g., Hoynes et al. (2016)) and use data on the fraction of all pregnant women who participated in WIC by race during a year in our study period (the only such data we could find), and treat that as the "first stage" for a long difference as pre-WIC measures of the WIC caseload are zero.

The requisite shares of women participating in WIC by race for that one year are reported in Figure 2.[33] These numbers from 1986 from Food Research and Action Center (1988) are the only information available, to the best of our knowledge, about WIC participation by race in this period. The WIC participation rate (the ratio of the number of Black participants who are women to the number of births to Black women in a given year) was 36% for Black pregnant women and 15% for White pregnant women.[34] These participation numbers may be an underestimate of the first stage because we are averaging participation across places with and without WIC. Using this information to scale our ITT thus provides an upper bound for the TOT. Our ITT estimates give an estimated TOT effect for low birth weight of

is the same, the bulk of the confidence intervals for the event study estimates by period after adoption are not crossed by the blue lines, which further validates our results for these outcomes. For Whites, post-WIC estimates are noisier and insignificant and overlap with the expected coefficients under a plausible violation of pre-trends. We further conduct Rambachan and Roth (2023) sensitivity analyses for potential violations of the parallel trend assumption (see figures A.8 and A.9). For Blacks, our event study estimates for the post-treatment periods are similar to those that allow for linear violations of the parallel trends, whereas confidence intervals widen slightly as we allow for non-linearity violations (not shown here). The results for Whites are more sensitive to these tests.

[33] The data sources for this analysis are listed in the table notes.

[34] If every county had WIC by 1986, then this would be a crude estimate of a long difference of the first stage, as no one was on WIC in 1968–1973, and 36% or 15% were on in 1986. However, these numbers are an average of participation in places with and without WIC; hence, the correct TOT adjustment might be to divide by something larger than these participation numbers because 36% average participation for Black mothers represents zero participation in some places and participation above 36% in others.

at most -0.9 percentage points for Black mom's babies (statistically significant) and at most -0.4 percentage points for White mom's babies (not statistically significant). These TOT effects correspond to a -6.4% of the control mean for Blacks and -6.1% for Whites.[35]

5.2 Distributional effects for birth weight vary by race

Our results to date suggest the possibility that WIC moved the entire birth weight distribution upward. We explore further heterogeneity in the effects of WIC on birth weight by looking at various cuts across the birth weight distribution, namely the probability that birth weight is below 1500, 1750, 2000, 2250, 2500, 2750, 3000, 3250, 3500, 3750, 4000, 4250, and 4500 grams. We present these results in Figure 5.

Regarding the distribution of birth weight by race, Panel (c) shows the share of the sample with a birth weight under each cut-point value for infants with Black mothers (in lavender color) and for infants with White mothers (in grey color). For example, around 13 percent of Black babies had a birth weight under 2500 grams (low birth weight), while only approximately 6 percent of White babies did. These means show that the White advantage in birth weight is present across the birth weight distribution.

Overall, we find that WIC shifts the entire birth distribution to the right. Panel (a) uses individual-level data to show the Sun and Abraham (2021) point estimates for these birth weight cuts. In addition to reducing the incidence of very low birth weight (< 1500 grams) and low birth weight (< 2500 grams), WIC decreased the share of Black births with birth weights under 1500 grams, 1750 grams, 2000 grams, 2250 grams, 3000 grams, and 3250 grams, as well as 3750 grams, 4000 grams, and 4250 grams. It led to a reduction in being below 4500 grams (and almost surely led to a small increase in the probability of being a high birth weight infant). For each of these measures, the White estimates are much smaller in magnitude. For the share of infants with birth weights under 1500 grams, 1750 grams, 2000 grams, and 2250 grams, the equality of estimates among races can likely be rejected at the 90% confidence level, and for some, we may be able to reject it at the 95% confidence level.

Panel (b) converts these estimates into percent effects by dividing the point estimates by the baseline means from Panel (c). The percent effects are also larger in magnitude for Blacks than for Whites, for the lower end of the birth weight distribution. This suggests that the racial heterogeneity in WIC effects is not simply driven by the differences in baseline

[35]For Black infants: $100 * (-0.3078/36) \approx 0.9pp$, corresponding $100 * \frac{100*(-0.3078/36)}{13.434} \approx -6.4\%$ and White infants: $100 * (-0.0585/15) \approx 0.4pp$, corresponding $100 * \frac{100*(-0.0585/15)}{6.432} \approx -6.1\%$.

mean outcomes for births to Black and White mothers.

5.3 Potential explanations underlying the race heterogeneity impacts

We have shown that the approximated TOT effects on low birth weight are likely larger in magnitude for Black births.[36] Why does WIC improve infant health outcomes for Black mothers more than for White mothers, even after adjusting for differences in the WIC participation rates by race?[37]

To better understand the race differences in WIC effects, we explored four general categories of possible explanations, in addition to the higher participation we already documented for Black mothers and the pre-WIC race differences in birth outcomes. These are as follows.

First, we examine county-level take-up data from 1980, predicting take-up rates as a function of exogenous (pre-WIC) maternal and county characteristics. Counties are then classified as being high in take-up or low in take-up, providing another measure of program reach. Differential take-up could explain our findings if Blacks and Whites lived in counties with different take-up rates or if Blacks participated at higher rates within the same locations. Second, we test whether the program is more effective for higher-need populations, either individuals with characteristics predicting worse outcomes or locations with, on average, worse pre-WIC birth outcomes (e.g., low birth weight) or more poverty. Third, we consider whether the location of medical infrastructure, proxied by hospitals per capita or the reach of the state's Medicaid programs at inception, explains the variation we find. Lastly, we consider two measures of racial animus. Following Almond et al. (2006), we identify states where a large share of Black, but not White, births occur outside of the hospitals, in the late 1960s, indicating limited hospital access for Black mothers. We also compared the results for counties in the Old Confederacy versus those outside of it. We explain our approach to consider each potential mechanism in detail below.

For mechanisms attached to individual-risk factors and location-based risk factors, we estimate models that fully interact the presence of the risk characteristic with all of the fixed

[36] Given that we do not have county-level data on eligibility or participation by race, we cannot do better than our relatively crude upper bound above for the magnitude of the TOT.

[37] We note that we have also estimated models for those of Other races, which combines pregnant women who are Asians, Alaska Natives, and Pacific Islanders, and found somewhat noisy results. The effect on birth weight for Other race mothers is the only marginally statistically significant result. It is quite large (a 21 grams improvement following WIC's introduction) relative to the estimated effects for Blacks (12 grams) and Whites (5 grams). These results are reported in Appendix Table A.8.

effects (controls). With this fully interacted model, the resulting average effects by subgroup can then be differenced and the differences tested to see if effects vary in a statistically significant fashion across sub-groups. The results for this analysis are shown in Figures 6, 7, and 8 for low birth weight.[38] In general, we present the point estimates and 95% confidence intervals for each subgroup in Panels (a) for Blacks and (b) for Whites, and the differences in point estimates and their 95% confidence interval by subgroups in Panels (c) for Blacks and (d) for Whites.

We first explore the role of take-up using the only county-level measure available. One of the Congressional reports had a list of WIC participating counties in 1980, along with take-up at the county level.[39] We predicted this take-up as a function of the baseline county controls and then estimated effects for counties with high (above-the-median) values of take-up versus low (at or below-the-median) values. The effects are slightly larger for Black mothers in counties with higher predicted take-up. This is shown in Figure 7, in the leftmost pair of bars in Panels (a) and (b) for Blacks and Whites, respectively (with the left bar representing the sample in counties with high predicted take-up, and the right bar representing those with low predicted take-up). Then, in Panels (c) and (d), the figures display the difference in the effects on low birth weight, along with the 95% confidence intervals indicating the statistical significance of these differences. The results show that places with higher predicted take-up had larger reductions in low birth weight for Black mothers but not for Whites.

Our second potential mechanism is either being an individual at higher risk of adverse birth outcomes or living in higher-risk places. We begin with individual risk factors. To start with, we show in Table 2 that WIC did not lead to large compositional changes in the share of births with these risk factors. For example, we found no effects of WIC access on the birth rate (Column 1), on the log of the number of births, on the share with teen mothers, on the share married, or on the share of mothers with low or high education. We find that county-level composition of parity is correlated with WIC exposure, but these differences are very small in magnitude.

Having ruled out substantive compositional changes due to WIC in the characteristics of mothers having children, we turn to regressions that stratify on the risk factors. These results are in Figure 6, with estimates for Blacks on the left (Panels (a) and (c)) and Whites on the right (Panels (b) and (d)). We compared the effect of WIC by race for teen to non-teen births, effects for births to unmarried mothers versus married mothers, births to more or less

[38]In this discussion, while we focus on the low birth weight outcome, we also present the analogous analysis on birth weight in the Appendix Figures A.10, A.11, and A.12.

[39]This report uses data from MIT researchers, including county participation divided by county eligibility to proxy WIC take-up and group the counties by take-up ranges.

highly educated women (in the restricted number of states that collect education status), and first-born versus later-born babies. We do not find strong evidence that the effects of WIC by race are particularly driven by one subsample versus the other (for example, the effects for less educated and more educated mothers are not statistically different from one another by race). The only exceptions is a larger reduction in low birth weight among White teen mothers compared to non-teen White mothers, and among mothers with less than a high school degree compared to more educated White mothers, a pattern not observed among Black mothers. This evidence suggests that these individual risk factors are not very likely to explain the differences in WIC impacts by race.

Next, we explored place-based risk factors such as county-level poverty concentration, racial composition, and urbanicity in 1970, and found these factors to be of limited relevance. We also considered an area measure of poor birth outcomes, predicting the baseline average of low birth weight (in 1970), with results not very conclusive (Figure 7).

Our third possible explanation was that racial differences might be explained by characteristics of the medical system/medical need in the county. Our measures here are hospitals per capita and the state share of Black participation in AFDC for women, but these results were also not very informative (see Figure 7's right most two estimates in Panels (a) and (b) and the right most one in Panels (c) and (d) and and 8).

Our fourth explanation was the role of proxies for structural racism. We do so by first comparing results within the states of the old Confederacy with the rest of the U.S., and do not find this to be informative. The second approach follows Almond et al. (2006) and uses the share of births taking place outside the hospital for Blacks in 1968 but not Whites as a measure of where hospitals were likely to be racially segregated. This set of states resembles the old Confederacy but does not include all the old Confederate states. Figure 8 shows these results; neither measure seems to explain the differences.

Overall, the differences in the effects of WIC by race are not explained by individual or place-based risk factors or structural racism as currently measured. The results may be explained by (1) baseline differences in outcomes by race (Figure A.1), therefore higher returns to investments for Black mothers relative to White mothers, (2) differences in who lives in places with larger or smaller predicted take-up, and (3) differences in the reach of WIC within location by race (as Blacks are on average more likely to be eligible).

5.4 Robustness checks

Our results for low birth weight and birth weight are highly robust to several alternative specifications as shown in Tables 3 and A.9, respectively. Column 1 reproduces our preferred specification. Next, we present specifications that, instead of using propensity score reweighting, only control for county and year-fixed effects (Column 2) and progressively add further War on Poverty program covariates (Column 3), county time-varying demographic, economic, and medical characteristics (Column 4), 1970 significant determinants interacted with linear county time trends (Column 5), and state-by-year fixed effects (Column 6). Lastly, in Column 7, we show the results using TWFE with propensity score reweighting. Overall, our results are reasonably stable across different specifications with varying sets of controls.[40]

Second, a potential concern is that there may be counties with a small fraction or zero Black births, which implies that the set of counties in the analyses by race may be different, and the county composition may explain the different results by race. To address that, we present estimations that limit the sample to a balanced set of counties that have at least 2 Black and White births each month and year of the second trimester (see Appendix Figures A.3 and A.4).[41] These tests show that the results are remarkably similar to the sample of all counties presented above.

Third, a related concern may be that the Panel of treated counties is unbalanced in event time, and changes in this composition of counties may be different by race. To assess this issue, Appendix Figures A.3 and A.4 present estimates with a balanced set of counties adopting with various leads and lags. Results are notably comparable to those from the full sample of counties.

Fourth, in Appendix Table A.10, we test whether our results are sensitive to alternative weights. For this, we show our preferred specification, which uses propensity score reweighting multiplied by the cell size of births. We then reestimate our results using only the propensity score reweighting without multiplying by births for all the counties in our sample (Column 2) and for a balanced set of counties with at least two Black and White births, multiplying by births (Column 3) or not doing so (Column 4). As expected, the results for the all-counties sample are somewhat sensitive to weighting by births because we effectively

[40]Across different specifications in Table 3, the point estimates on low birth weight range between 0.22 to 0.31 percentage points for Blacks and between 0.013 and 0.05 percentage points for Whites. Appendix Table A.9 shows that birth weight point estimates range between 5 to 12 grams for Blacks and between 1.8 to 5 grams for Whites.

[41]Note that this sample restriction limits the number of counties in the estimation to 455.

assume that small and large counties are equally weighted in the estimation. However, when we restrict to a balanced set of counties, our results are similar when we weight or do not weight our estimates by the number of births, in addition to the pscore.

Fifth, we have performed additional robustness checks in Appendix Table A.11 that include: Using individual-level data instead of cell-level data (which should be identical with weighted regression and linear regression), alternative assumptions when computing gestational age, and alternative assignment of WIC timing (month and year of the third trimester). Our results using individual-level data in Columns 1 and 6 are identical to the cell-level results weighted by births (Table 1, as expected). In Columns 2, 3, 4 for Black births and 7, 8 and 9 for White births, we perform several robustness checks using alternative assignments of gestational age. We do this because, for a subset of the sample, gestational age is missing. In our preferred estimation, we assign 30.5 days to a month and 40 weeks to records missing gestational age to construct gestational age for the full sample following Currie and Schwandt (2013). However, our results are not sensitive to using 30 days as the average number of days in a month (Column 2), assigning 39 weeks for the missing gestational age (Column 3), or keeping the gestational age as missing in our sample (Column 4). Lastly, our results are very similar if we assign WIC exposure according to the month-year of the start of the third trimester (Columns 5 and 10).

Additionally, we performed robustness checks that collapse and assign exposure at the year level instead of the month-year (Appendix Table A.12). In particular, we assign exposure based on the year of the second trimester, the year of the third trimester, and the year of birth. These estimates are quite comparable to those from our preferred specification, but smaller as we are ignoring more granular variation.

Sixth, we check if our estimates are sensitive to trimming the propensity score to address potential concerns about counties with a very high propensity of having WIC affecting the results, as well as to controlling for the presence of the Commodity Supplemental Food Program (CSFP), which provided nutritional assistance similar to WIC to low-income families. We present these results in Appendix Table A.13. In Columns 1 and 4, we show our preferred specification, not trimming the propensity score and controlling for CSFP presence. In Columns 2 and 5, we show that the results are not sensitive to trimming the propensity score by dropping counties with more than a 0.99 probability of having WIC. In Columns 3 and 6, we present our results without the presence of CSFP, which show that the point estimates and the standard errors are slightly smaller but economically meaningful.

Seventh, our compilation of WIC rollout includes programs that served special populations only, including migrants and Native Americans. Appendix Table A.14 shows that our findings on the effect of WIC on birth outcomes are not driven by these places and are highly robust to excluding those counties (see Columns 2 and 4).

6 Discussion

We consider the magnitude of the effects we have uncovered in this section, placing them in the context of previous literature, particularly Hoynes et al. (2011) (HPS) and other work on another War on Poverty program, Food Stamps.

Since we build upon the HPS WIC rollout study, we present estimations that replicate their specifications. There are three reasons our results may differ from theirs: First, it could be due to our adding additional states and years, and second, it could be because we could uncover additional places where WIC rolled out in counties they included as controls with the additional archival data on counties which were served by existing WIC local agencies. Table A.2 shows that we added nearly 485 new counties to their data, with many adopting WIC in the years they could not include in their analysis (1976, 1977, and 1980), but also because we can show some of them adopted earlier (or less often, later). If we were simply adding to our treatment group more counties/years that adopted WIC to the analysis with the same counties and years, we would expect our point estimates to be larger in magnitude, as some treated locations would have been included in the control group. It is also possible that their sample, which excludes some years and states, may have led to bias due to using a time-varying set of cross-sectional units (states) in some periods relative to others. Finally, we use exposure in the year of the beginning of the second trimester as our treatment variable, which is likely conservative for measuring when most women got onto WIC.

Our replication does the analysis using their sample construction steps and years.[42] These results are presented in Appendix Table A.15. HPS use TWFE models and present two specifications. The first specification includes county and year fixed effects, time-varying per capita county transfers, per capita income, Food Stamp exposure, and 1970 county variables interacted with a linear time trend. The second specification adds state-by-year fixed effects. Columns 1 and 2 use their original rollout, and all their counties, and show that estimates are similar to those reported in their paper. We corroborate their finding of an increase

[42] Hoynes et al. (2011) estimate their analysis using cell-level data collapsed to the "county and year of the third-trimester" level for 1971–1975 and 1978–1982 and TWFE specifications. The reason they exclude 1976 and 1977 is their lack of primary sources for WIC rollout during those years. In addition, they limit their analysis to the set of counties that have WIC in place by 1979, whereas we add 1980 rollout. WIC exposure for HPS is assigned as of the year of the third trimester. Their sample also excludes cells that contain fewer than 25 births in 1971–1975 and 1978–1982.

in birth weight but no effect on the probability of low birth weight using their sample and specification. Columns 3 and 4 use their rollout but also exclude counties in Virginia and Hawaii (as we do in our analysis) and show similar effects. In Columns 5 and 6, we show that their results are robust when using our rollout variable applied to their sample of counties and specification. Since we added WIC rollout information for a sizable number of counties, Columns 7 and 8 use their years of data and specifications for our augmented sample of counties in those years. The point estimates for birth weight and low birth weight using our sample and coding of WIC are larger in magnitude. For example, for birth weight, our point estimates are larger by about 17 percent (comparing Column 3 to Column 7) or about 12 percent (comparing Column 4 to Column 8), highlighting the importance of incorporating the additional rollout information.[43]

Finally, we compare our effects to some from other literature. The most obvious comparison is to Almond et al. (2011), which estimates the effect of Food Stamp rollout on birth outcomes. We report our and their treatment effects on the treated in Table 4. Our overall reduced form effects for infants of Black mothers were about twice as big as theirs for low birth weight (Panel A, column 1) and about 3.4 times as big for birth weight (Panel A, column 4). Turning to TOT estimates for each paper (Columns 3 and 6), we see that the TOT for Black infants for FS rollout was only a 0.37 percentage point reduction in being low birth weight, while the TOT for infants of Black mothers for WIC rollout was a 0.9 percentage point reduction in being low birth weight. Furthermore, the programs had similar reach, with our administrative data suggesting 36% of Black pregnant women and 15% of White pregnant women participating in WIC in 1986, while HPS estimate that 41% of Black pregnant women and 13% of White pregnant women participated in Food Stamps in 1980.

7 Conclusion

The WIC program, which first distributed benefits in 1974, provides low-income, nutritionally at-risk women with vouchers for foods that are rich in nutrients of concern during pregnancy; such as iron, calcium, and vitamins A and D. In this paper, we have comprehensively explored the effect of the introduction of the WIC program at the county level on birth outcomes from 1968–1980. We do so using greatly expanded data on the rollout of the program from the National Archives and other sources, and using approaches robust to concerns about treatment effect heterogeneity in average treatment effects over time and

[43] Comparing Column 3 and 7 for birth weight, we find $100 * \frac{(3.0458-2.5933)}{2.5933} = 17.45\%$. Comparing Column 4 and Column 8 for birth weight, we find $100 * \frac{(2.2554-2.0080)}{2.008} = 12.32\%$.

place, and robust to contamination in event-study approaches.

We use Sun and Abraham (2021)'s approach with never-treated counties as the control group (counties adopting WIC after 1980 in our context). As counties that implemented WIC by 1980 may be different than those that did not, we adjust for geographic differences in pre-program characteristics, including socio-demographic characteristics, presence of other food and health programs, and births' characteristics, with inverse propensity score weights from predicting county WIC adoption. We find robust evidence that WIC moved the entire birth weight distribution up for Black mothers, with economically meaningful and statistically significant reductions in low birth weight (2.3 percent) and being below a number of birth weight thresholds from 1500 grams through 4500 grams. There are statistically significant effects for White mothers, though smaller in magnitude, at birth weight thresholds below 2750, 3000, 3250, and 3500 grams. When translated to percent effects (divided by the baseline mean share of births under these thresholds), the effects are still larger in magnitude for Black mothers, and different from those for Whites for thresholds below 2250 grams.

We have also explored a host of other outcomes, such as being small for gestational age, being preterm, birth weight, high birth weight, being large for gestational age, and one measure of use of the medical system, getting adequate prenatal care (Kessner Index). For Blacks, we find a decline in the likelihood of being very low birth weight (-5.6 percent) and small for gestational age (-4.5 percent), as well as an increase in large for gestational age (10.9 percent), and in average birth weight (11.7 grams), all of which are significant at the 5 percent level. For Whites, we find an increase in average birth weight of 5.1 grams that is significant at the 5 percent level (increase of 0.15 percent).

Our results are robust to sensible changes in specification, and we find no evidence that there were important compositional changes in the characteristics of mothers giving birth or in birth rates. We explored and ruled out concerns about pre-trends driving low birth weight, and small for gestational age for Black mothers, using the approaches of Roth (2022) and Rambachan and Roth (2023). For White mothers, several more results may be explained by pre-trends.

We are also able to uncover administrative data on WIC participation by race for the full population of women in 1986 to get a back-of-the-envelope measure of the effect of WIC implementation on use of the program as a "first stage." WIC eligibility was extremely high for blacks. Nearly 75 percent of Black women were eligible for WIC, and 36 percent participated in WIC. White eligibility and participation were also substantial, though lower than for Black women, with 36 percent of White mothers being eligible and 15 percent participating. If these were the appropriate first-stage estimates for WIC participation, they would imply treatment-on-the-treated (TOT) effects corresponding to a decline of low birth weight of 6.4 percent for Blacks and 6.1 percent for Whites. The analogous effects on being small for gestational age are a decline of 12.5 percent from WIC implementation for Black mothers, and 7.4 percent for White mothers. These are clearly economically meaningful.

Next, we explored a range of possible explanations for the larger impacts observed among Black mothers compared to White mothers, estimating models pooling the groups and formally testing differences across subgroups. First, we consider the reach of the program, using predicted take-up of WIC at the county level from USDA participation, coupled with estimates of eligibility at the county level. If WIC had higher predicted take-up where there are more Black mothers, this could explain the racial differences we see. Then, we consider the possibility that effects might be larger where the risk of negative birth outcomes is higher. To explore this, we examine subgroups based on individual-level risk factors (maternal age and education, marital status, and parity) as well as area-level risk measures from 1970 (county-level predicted low birth weight rate, poverty rate, urban/rural status, and the share of non-White residents in the county). We also consider whether differences are explained by the level of hospitals per capita, a measure of the adequacy of the health care system in the county, as well as by the magnitude of the state-level reach of the Medicaid program at implementation, proxied by the non-White female participation rate in AFDC, as suggested by Goodman-Bacon (2021). Our last two measures are both state-level proxies for racial animus. First, we compare states that were part of the old Confederacy with those that were not. Second, we consider a proxy suggested by Almond et al. (2006), whether the state had a large share of black but not white births occurring outside the hospital in 1968. In sum, we find that the most relevant features where there is a significant difference in the effects of WIC on low birth rates across race groups are for places with higher than median predicted WIC take-up (which benefit more from WIC). By contrast, for Whites, effects are larger in magnitude for teens and for mothers with less than a high school degree.

Overall, we have found striking evidence that the effects of WIC implementation for Blacks were important and moved the entire black birth weight distribution upward. Effects were larger where the program reached more eligible women. By contrast, effects for Whites, while smaller in magnitude, were larger for teen mothers and for mothers with less than a high school degree. Effects for WIC are larger in magnitude than those for the Food Stamp Program, the rollout of which has been shown to lead to improvements in life outcomes in adulthood. This shows the success of this program in its early years, and calls for work looking at the long-run effects of the WIC program.

References

Almond, D., K. Y. Chay, and M. Greenstone (2006). Civil rights, the War on Poverty, and black-white convergence in infant mortality in the rural South and Mississippi.

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

Almond, D., H. W. Hoynes, and D. W. Schanzenbach (2011). Inside the War on Poverty: The impact of Food Stamps on birth outcomes. *The Review of Economics and Statistics 93*(2), 387–403.

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

Bailey, M. J., H. W. Hoynes, M. Rossin-Slater, and R. Walker (2022). Is the social safety net a long-term investment? Large-scale evidence from the Food Stamps Program. *The Review of Economic Studies*.

Bendick Jr., M., T. Campbell, D. L. Bawden, and M. Jones (1976). Efficiency and effectiveness in the W.I.C. program delivery system. Technical report, USDA Miscelleous Publication 1338, FNS, USDA.

Benjamini, Y., A. Krieger, and D. Yekutieli (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika 93*(3), 491—-507.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. *The Quarterly Journal of Economics 122*(1), 409–439.

Callaway, B. and P. H. Sant'Anna (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics 225*, 200–230.

Chambers, B. D., J. T. Erausquin, A. E. Tanner, T. R. Nichols, and S. Brown-Jeffy (2018). Testing the Association Between Traditional and Novel Indicators of County-Level Structural Racism and Birth Outcomes among Black and White Women. *Journal of Racial and Ethnic Health Disparities 5*(5), 966–977.

Cunningham, J. and A. Goodman-Bacon (2024). Changes in Family Structure and Welfare Participation since the 1960s: The Role of Legal Services. *American Economic Journal: Applied Economics 17*(1), 369–401.

Currie, J. (2011). Inequality at birth: Some causes and consequences. *The American Economic Review 101*(3), 1–22.

Currie, J. and I. Rajani (2015). Within-mother Estimates of the Effects of WIC on Birth Outcomes in New York City. *Economic Inquiry 53*(4), 1691–1701.

Currie, J. and M. Rossin-Slater (2020). Does the WIC program promote equality of opportunity in early life?

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

de Chaisemartin, C. and X. D'Haultfœuille (2020, September). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–2996.

East, C. N., S. Miller, M. Page, and L. R. Wherry (2023, January). Multigenerational Impacts of Childhood Access to the Safety Net: Early Life Exposure to Medicaid and the Next Generation's Health. *American Economic Review 113*(1), 98–135.

Figlio, D., S. Hamersma, and J. Roth (2009). Does prenatal WIC participation improve birth outcomes? New evidence from Florida. *Journal of Public Economics 93*, 235–245.

Food Research and Action Center (1988). WIC facts: National and state profiles of the Special Supplemental Food Program for Women, Infants, and Children. Technical report, Food Research and Action Center.

Goodman-Bacon, A. (2018, February). Public Insurance and Mortality: Evidence from Medicaid Implementation. *Journal of Political Economy 126*(1), 216–262.

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

Goodman-Bacon, A. (2021, August). The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health, and Labor Market Outcomes. *American Economic Review 111*(8), 2550–2593.

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics 6*(2), 65–70. Publisher: [Board of the Foundation of the Scandinavian Journal of Statistics, Wiley].

Hoynes, H., M. Page, and A. H. Stevens (2011). Can targeted transfers improve birth outcomes? Evidence from the introduction of the WIC program. *Journal of Public Economics 95*(7–8), 813–827.

Hoynes, H., D. W. Schanzenbach, and D. Almond (2016, April). Long-Run Impacts of Childhood Access to the Safety Net. *American Economic Review 106*(4), 903–934.

Hoynes, H. W. and D. W. Schanzenbach (2009). Consumption responses to in-kind transfers: Evidence from the introduction of the Food Stamp Program. *American Economic Journal. Applied Economics 1*(4), 109.

Kennedy-Moulton, K., S. Miller, P. Persson, M. Rossin-Slater, L. Wherry, and G. Aldana (2022). Maternal and Infant Health Inequality: New Evidence from Linked Administrative Data. Technical Report w30693, National Bureau of Economic Research, Cambridge, MA.

Kose, E., S. O'Keefe, and M. Rosales-Rueda (2024). Does the delivery of primary health care improve birth outcomes? Evidence from the rollout of Community Health Centers. *Journal of Human Resources*.

Kramer, M. S. (1987a). Determinants of low birth weight: Methodological assessment and meta-analysis. *Bulletin of the World Health Organization 65*(5), 663.

Kramer, M. S. (1987b). Intrauterine growth and gestational duration determinants. *Pediatrics 80*(4), 502–511.

Ludwig, J. and D. L. Miller (2007). Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design*. *The Quarterly Journal of Economics 122*(1), 159–208.

Malinovskaya, A. (2023). *Long-Term Health Effects of the WIC program and policy multiplier effects of large scale public programs in the US*. Ph. D. thesis, Cornell Applied Economics and Management.

Manchester, A. (1976). Relationships of food costs, expenditures, and prices, 1960–1974. Technical report, Economic Research Service, USDA, Agricultural Economic Report 329.

Marshall, N. E., B. Abrams, L. A. Barbour, P. Catalano, P. Christian, J. E. Friedman, W. W. Hay Jr, T. L. Hernandez, N. F. Krebs, E. Oken, et al. (2022). The importance of nutrition in pregnancy and lactation: Lifelong consequences. *American Journal of Obstetrics and Gynecology 226*(5), 607–632.

National Research Council and Division of Behavioral and Social Sciences and Committee on Child Development Research and Public Policy and Panel for the Study of the Policy Formation Process (1982). *Making policies for children: A study of the federal process*. Washington, DC: National Academies Press.

Owens-Young, J. and C. N. Bell (2020). Structural Racial Inequities in Socioeconomic Status, Urban-Rural Classification, and Infant Mortality in US Counties. *Ethnicity & Disease 30*(3), 389–398.

Rambachan, A. and J. Roth (2023, September). A More Credible Approach to Parallel Trends. *Review of Economic Studies 90*(5), 2555–2591. Publisher: Oxford University Press (OUP).

Rauscher, E. and D. E. Rangel (2020). Rising inequality of infant health in the U.S. *SSM–Population Health 12*, 100698.

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

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

Roth, J. (2022, September). Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends. *American Economic Review: Insights 4*(3), 305–322.

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

Shapira, N. (2008). Prenatal nutrition: A critical window of opportunity for mother and child. *Women's Health 4*(6), 639–656.

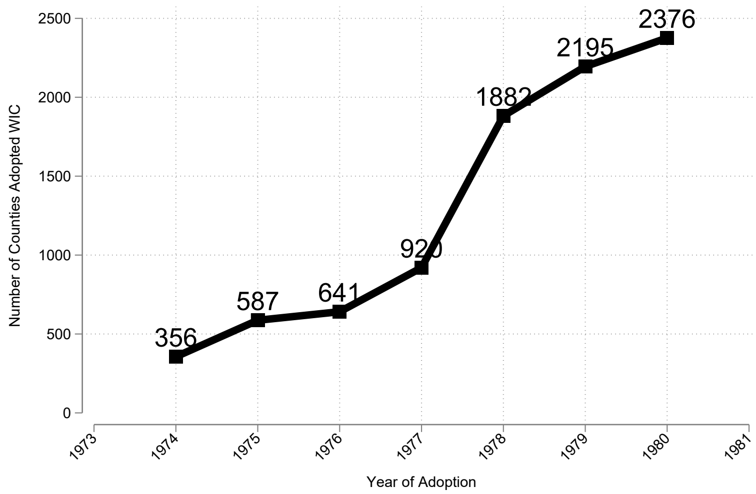
Singh, G. K. and S. M. Yu (2019). Infant Mortality in the United States, 1915-2017: Large Social Inequalities have Persisted for Over a Century. *International Journal of Maternal and Child Health and AIDS (IJMA) 8*(1), 19–31.

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

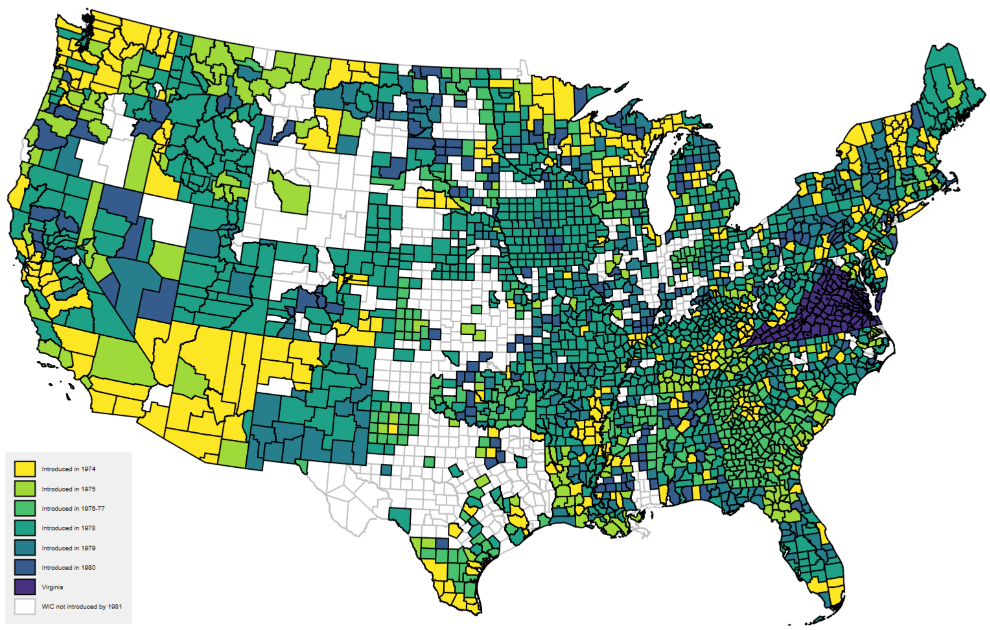
Vu, H., T. L. Green, and L. E. Swan (2024). Born on the wrong side of the tracks: Exploring the causal effects of segregation on infant health. *Journal of Health Economics 95*, 102876.

Figures

Figure 1: Geographic Variation and Timing of WIC Adoption



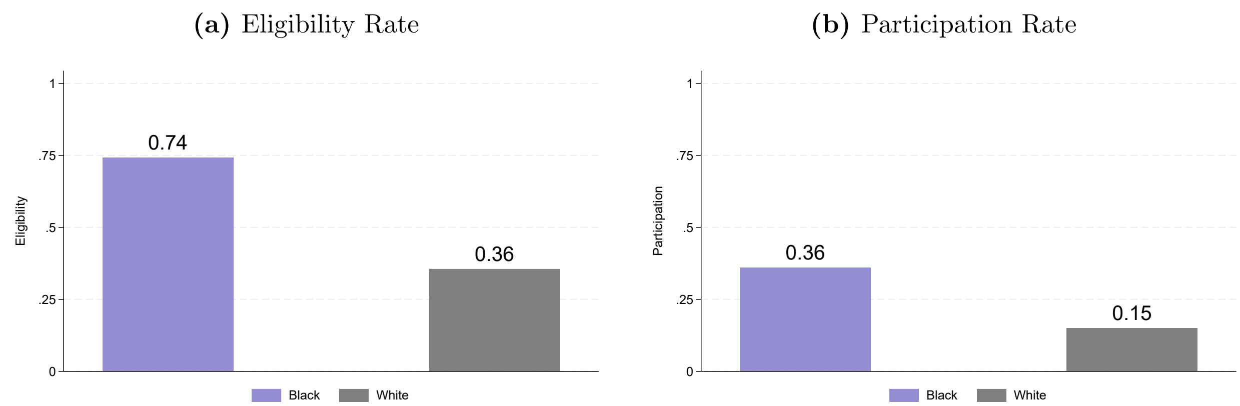
(a) Number of Counties Adopting WIC over Time



(b) Geographic Variation in WIC Adoption Timing

Notes: Our compilation of WIC adoption using the sources described in Data Appendix A. We do not include Virginia, Alaska, Hawaii, or other US territories in our study. After the restriction, there are a total of 2992 counties in the US. By 1980, around 79% (= $100 * 2376/2992$) of the counties had WIC.

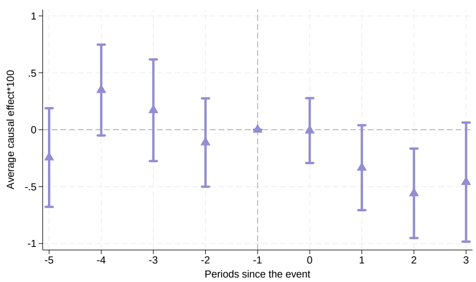
Figure 2: WIC Eligibility and Participation Are Higher among Blacks than Whites



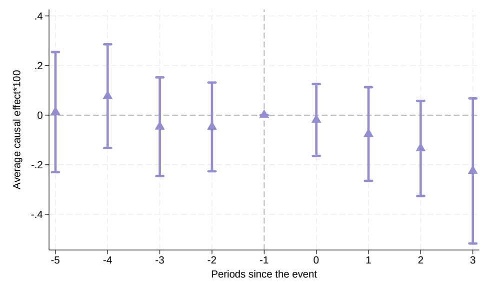
Notes: Eligibility calculations use measures of participation in WIC by race (Black, White, Asian/Pacific Islander, Native American, and Hispanic), and also broken down into women, infants, and children for Fiscal Year 1986 from (Food Research and Action Center, 1988). First, we combine the ITO counts with the counts for states where the ITO headquarters is located. We construct the estimates of participation by pregnant women by race as the product of the total number of Black or White participants times the share of participants in the state (note that the share of lactating/postpartum participants is quite low across our time period). We combine Hispanics with Whites to construct a "White" measure similar to that in our Natality data. Eligibility rates for White and Black Women are shown in Panel (a) and use the above participation measure as the numerator and a measure of eligibility for WIC from the March CPS for 1988–1990 for the denominator. The number of eligible pregnant women by race is given by the weighted sum of mothers of the relevant race living with a 0-year-old in the March CPS (and thus likely pregnant during the previous year) whose CPS family (pooling related subfamilies with the family of the household heads) had total income under the upper-income limit at the time from (Food Research and Action Center, 1988) or who were on AFDC or getting food stamps or free or reduced school lunch or were on Medicaid (all ways to be eligible). This was summed using the person weight for the 0-year-olds. Participation shares (Panel (b)) use the participants as the numerator and the weighted sum of births in the state from the Natality data in the denominator. The CPS data do not report Medicaid use before 1988.

Figure 3: Event Study: The Effect of WIC on Main Birth Outcomes—Black Mothers

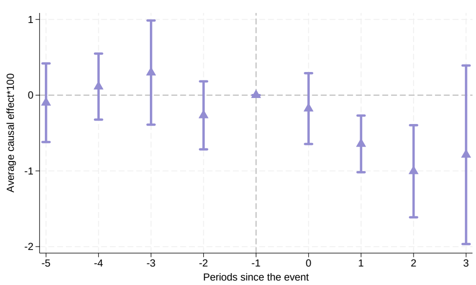
(a) Low Birth Weight



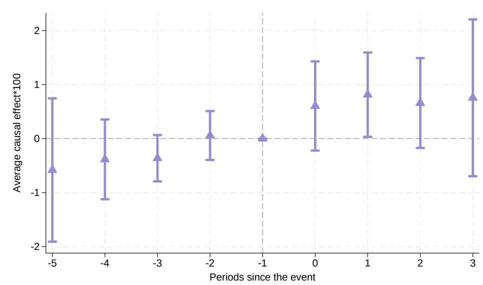
(b) Very Low Birth Weight



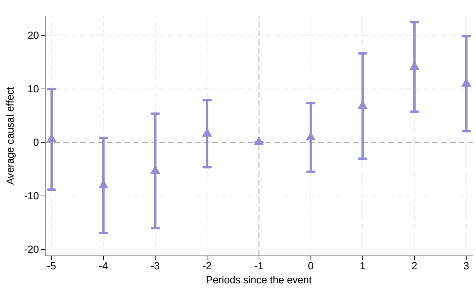
(c) Small for Gestational Age



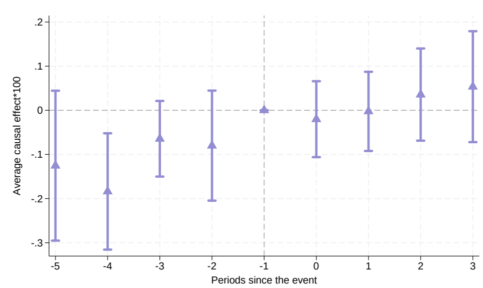
(d) Preterm



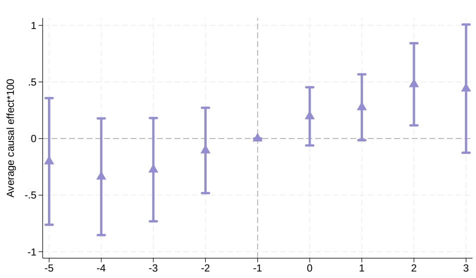
(e) Birth Weight



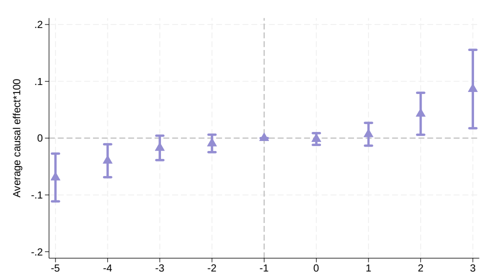
(f) High Birth Weight



(g) Large for Gestational Age



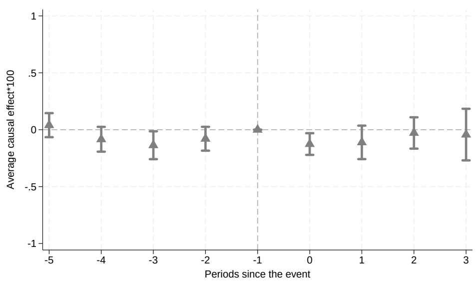
(h) Kessner Index-Adequate Prenatal Care



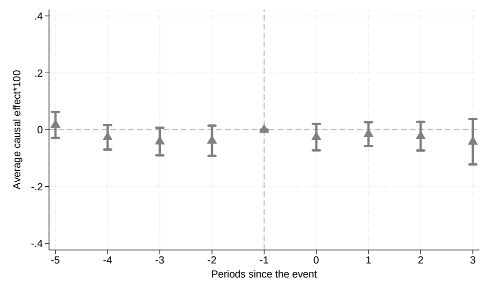
Notes: Figures report the estimated coefficients and their 95% confidence intervals using the Sun and Abraham (2021) method for Black mothers. All the dependent variables are multiplied by 100 (except for birth weight, which is measured in grams). Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells. We binned relative time periods more than 6 years before and more than 4 years after WIC adoption. Standard errors are clustered at the county level.

Figure 4: Event Study: The Effect of WIC on Main Birth Outcomes—White Mothers

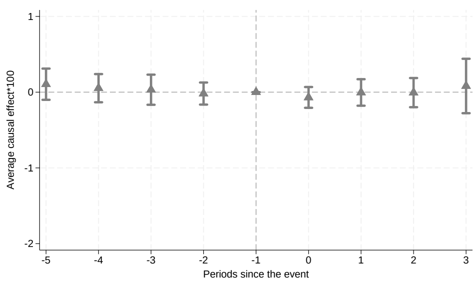
(a) Low Birth Weight



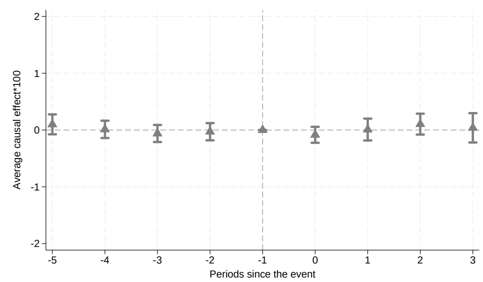
(b) Very Low Birth Weight



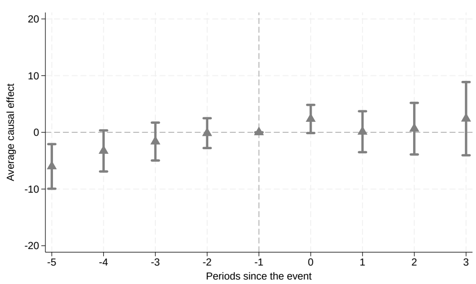
(c) Small for Gestational Age



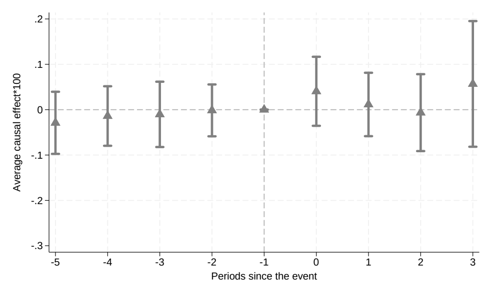
(d) Preterm



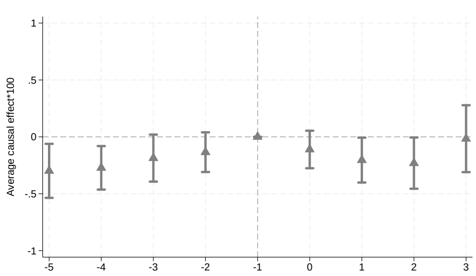
(e) Birth Weight



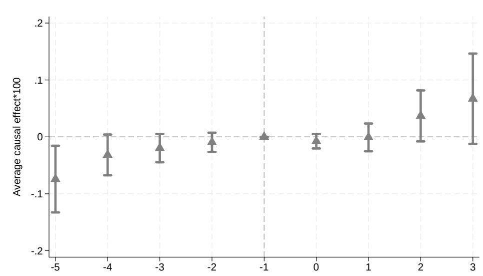
(f) High Birth Weight



(g) Large for Gestational Age



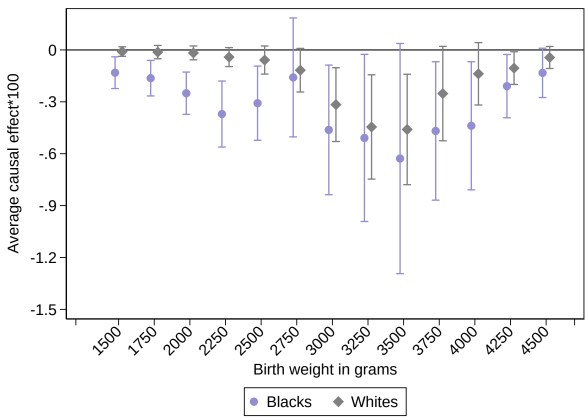
(h) Kessner Index-Adequate Prenatal Care



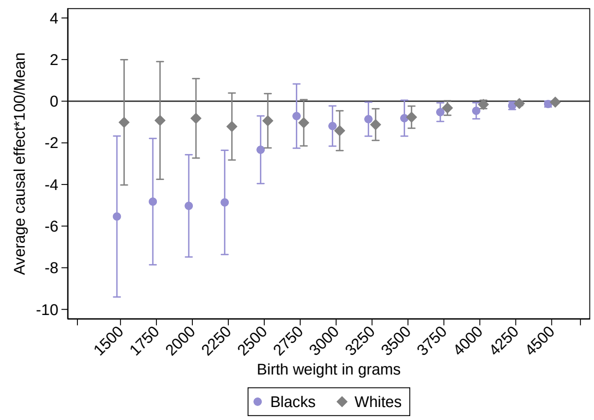
Notes: Figures report the estimated coefficients and their 95% confidence intervals using the Sun and Abraham (2021) method for White mothers. All the dependent variables are multiplied by 100 (except for birth weight, which is measured in grams). Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1 - \hat{p})$ times the number of births in the cell for the control group. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells. We binned relative time periods more than 6 years before and more than 4 years after WIC adoption. Standard errors are clustered at the county level.

Figure 5: The Effects of WIC on the Probability that Birth Weight is Less than Various Cut Points

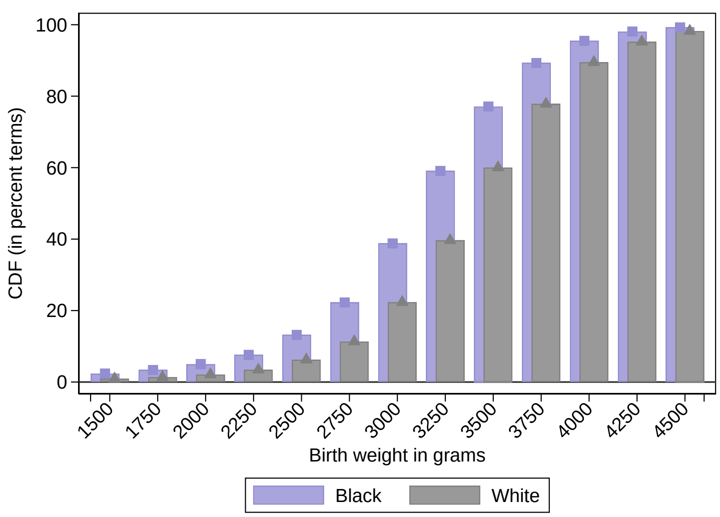
(a) Estimated Effect



(b) Percent Effect



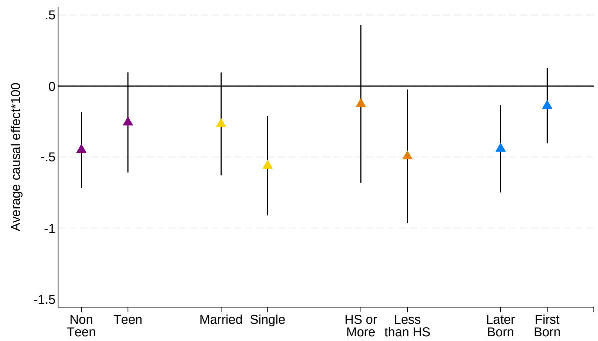
(c) Average Probability of Birth Weight < Each Birth Weight Threshold



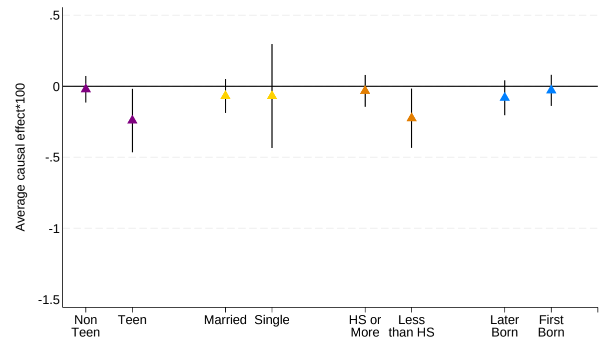
Notes: We estimate the effect of WIC on 100 times the probability that birth weight is below various gram thresholds, with cutpoints ranging from 1,500 to 4,500 grams in 250-gram increments. We plot the point estimates and the 95% confidence intervals and "% Impacts [coef./mean]" separately using the Sun and Abraham (2021) method. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. Standard errors are clustered at the county level.

Figure 6: Heterogeneity Analyses: Differential Effects of WIC by Individual Risk Factors—Low Birth Weight

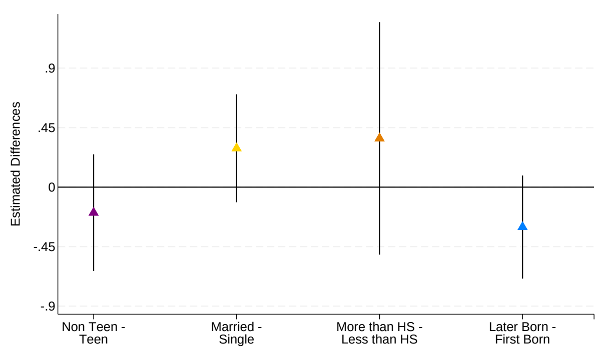
(a) Low Birth Weight—Black



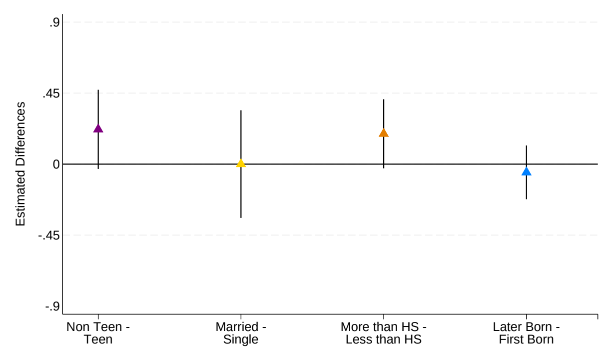
(b) Low Birth Weight—White



(c) Differences (more-less advantaged)—Black



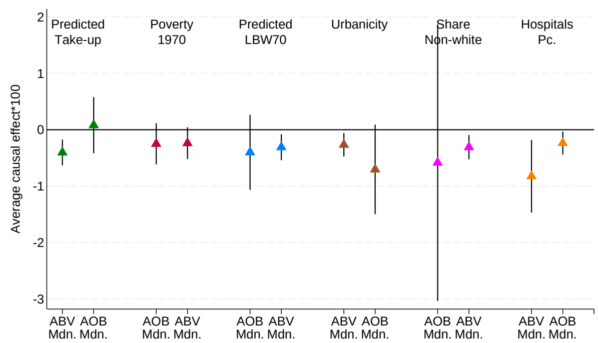
(d) Differences (more-less advantaged)—White



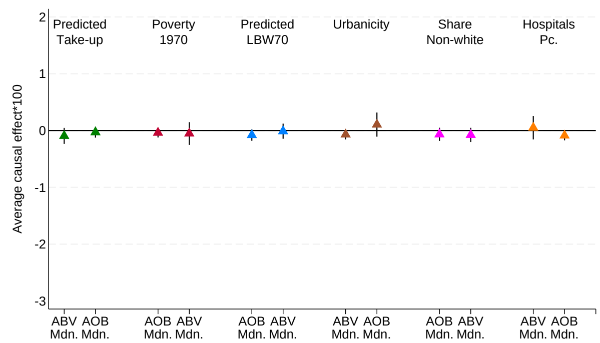
Notes: This analysis reports the estimated coefficients and the 95% confidence intervals, separately for each subgroup using the Sun and Abraham (2021) method. Each color represents a risk factor and corresponds to a separate regression. Each risk factor is defined based on the mothers' characteristics reported in the Vital Statistics. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. Panels A and B present estimates and 95% confidence intervals from fully saturated regressions, where each risk factor is fully interacted with all of the fixed effects within each race group. Panels C and D present the corresponding estimated difference between each factor group and its confidence interval. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by county, race, month, and year of the second trimester, parity, age category, gender, marital status, and education (finer cells than in our main analysis sample). Standard errors are clustered at the county level.

Figure 7: Heterogeneity Analyses: Differential Effects of WIC by County Risk Factors—Low Birth Weight

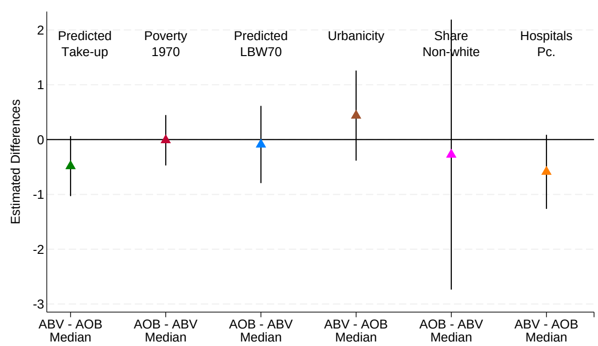
(a) Low Birth Weight—County Risk—Black



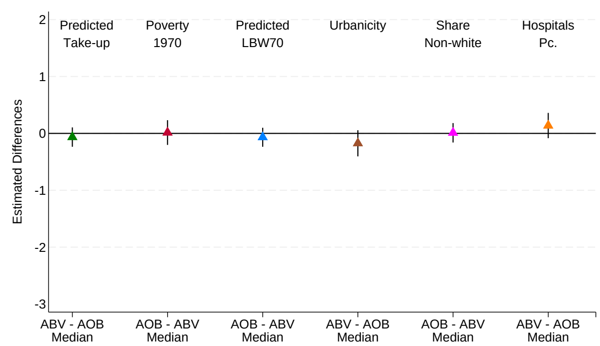
(b) Low Birth Weight—County Risk—White



(c) Differences across County Risk (more-less advantaged)—Black

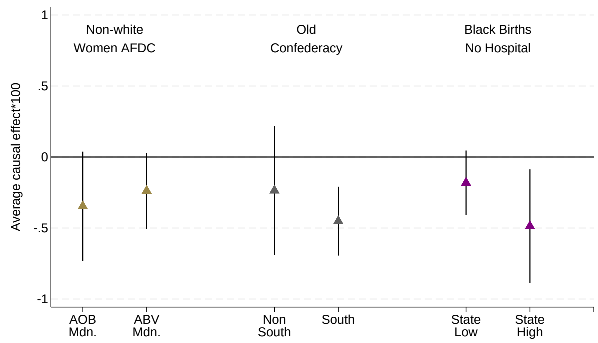


(d) Differences across County Risk (more-less advantaged)—White

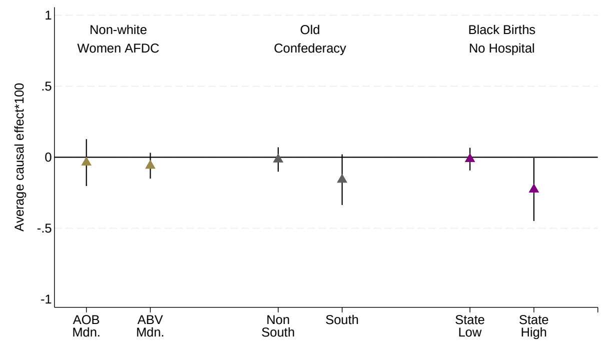


Notes: This analysis reports the estimated coefficients and the 95% confidence intervals, separately for each subgroup using the Sun and Abraham (2021) method. Each color represents a risk factor and corresponds to a separate regression. Each risk factor is defined based on county characteristics compiled using various data sources, as above the median value or at or below the median value of the county characteristics, with the unweighted median calculated using all counties after dropping VA, AK, and HI. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. Panels A and B present estimates and 95% confidence intervals from fully saturated regressions, where each risk factor is fully interacted with all of the fixed effects within each race group. Panels C and D present the corresponding estimated difference between each factor group and its confidence interval (i.e., counties at or below (AOB) the median of 1970 poverty rate and counties above the median (ABV)). We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by county, race, month, and year of the second trimester, parity, age category, gender, marital status, and education (finer cells than in our main analysis sample). Standard errors are clustered at the county level.

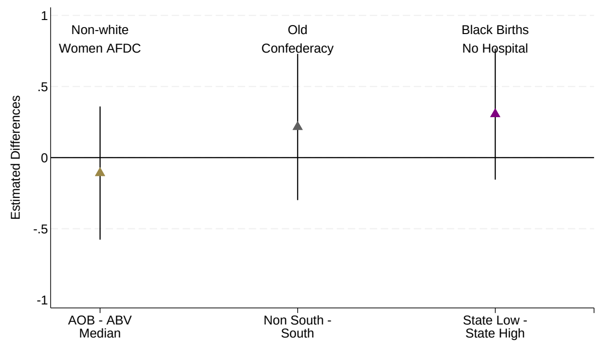
Figure 8: Heterogeneity Analyses—State Risk Factors—Low Birth Weight



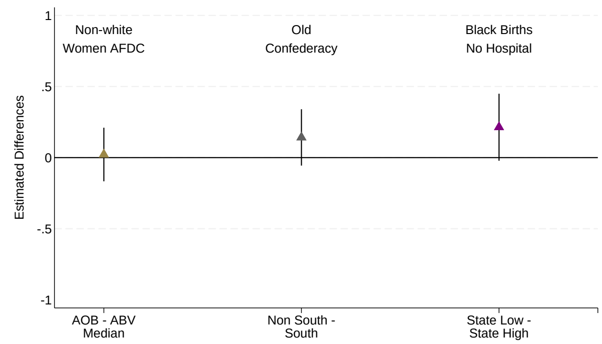
(a) Low Birth Weight—State Risk—Black



(b) Low Birth Weight—State Risk—White



(c) Differences across State Risk (more-less advantaged)—Black



(d) Differences across State Risk (more-less advantaged)—White

Notes: This analysis reports the estimated coefficients and the 95% confidence intervals, separately for each subgroup using the Sun and Abraham (2021) method. Each color represents a risk factor and corresponds to a separate regression. Each risk factor is defined based on state characteristics. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. Panels A and B present estimates and 95% confidence intervals from fully saturated regressions, where each risk factor is fully interacted with all of the fixed effects within each race group. Panels C and D present the corresponding estimated difference between each factor group and its confidence interval (i.e., Non-Old Confederacy states and old Confederacy states). Non-white women in AFDC counts are from Goodman-Bacon (2018). States of the former Confederacy include: Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia. The group of states with a high share of Black births not occurring in the hospital is based on information reported in the Vital Statistics in 1968, for states reporting more than 5% of Black but not White births occurred outside the hospital. These states include: Alabama, Arkansas, Florida, Georgia, Mississippi, North Carolina, South Carolina, Texas, and Virginia. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by county, race, month, and year of the second trimester, parity, age category, gender, marital status, and education (finer cells than in our main analysis sample). Standard errors are clustered at the county level.

Tables

Table 1: The Effect of WIC on Main Birth Outcomes

	Low Birth Weight	Very Low Birth Weight	Small for Gest. Age	Preterm	Birth Weight	High Birth Weight	Large for Gest. Age	Kessner Index Adequate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Black Mothers								
WIC-month year-trim2	-0.3078***	-0.1315***	-0.6518**	1.0531	11.7089***	0.1306*	0.6588***	0.0315*
	(0.1095)	(0.0469)	(0.3130)	(0.6836)	(4.0157)	(0.0727)	(0.2486)	(0.0171)
Control Y-mean	13.434	2.362	14.437	17.880	3080.080	0.734	6.057	0.185
% Effect (Coef/Control Mean)	-2.291	-5.570	-4.515	5.890	0.380	17.806	10.876	17.021
Adj. R-Squared	0.028	0.014	0.044	0.069	0.079	0.031	0.046	0.606
Obs	189417	189417	142988	142988	189417	189417	142988	147784
B: White Mothers								
WIC-month year-trim2	-0.0585	-0.0093	-0.0958	0.0238	5.1234***	0.0398	0.1522*	0.0107
	(0.0415)	(0.0141)	(0.0929)	(0.0794)	(1.8842)	(0.0328)	(0.0877)	(0.0211)
Control Y-mean	6.432	0.933	8.612	8.429	3339.085	1.663	10.092	0.282
% Effect (Coef/Control Mean)	-0.909	-0.997	-1.112	0.282	0.153	2.390	1.508	3.778
Adj. R-Squared	0.053	0.008	0.070	0.043	0.269	0.046	0.076	0.709
Obs	440037	440037	378283	378283	440037	440037	378283	346420
C,Year FE	Y	Y	Y	Y	Y	Y	Y	Y
1970 controls-pscore weighting	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Table reports the estimated coefficients, control mean (in percentage point terms (probability*100) for all the variables except for birth weight, which is in grams), and standard errors using the Sun and Abraham (2021) method, controlling for county and year of the second trimester. Panel A presents results for Black mothers, and Panel B for White mothers. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells. Standard errors are clustered at the county level. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

Table 2: The Effect of WIC on the Composition of Births

	Birth Rate	Ln(Births)	Parity	Mother's Age	% Teen Mother	% Married	%< HS Graduate	% HS Graduate	% > HS Graduate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
WIC-month year-trim2	-0.0084	-0.0048	0.0197*	0.0495	0.1871	-0.8659	0.7258	-0.5129	-0.2130
	(0.0490)	(0.0161)	(0.0102)	(0.0301)	(0.2330)	(0.7295)	(0.4981)	(0.4715)	(0.2843)
Y-mean	9.124	5.572	2.102	24.389	18.037	87.726	27.791	46.889	25.320
% Effect (Coef/Control Mean)	-0.092	-0.087	0.937	0.203	1.037	-0.987	2.612	-1.094	-0.841
Adj. R-Squared	0.862	0.983	0.449	0.604	0.489	0.796	0.749	0.474	0.729
Obs	440953	440953	441889	441889	441889	310420	294660	294660	294660
C,Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
1970 controls -pscore weighting	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Table reports the estimated coefficients and standard errors using the Sun and Abraham (2021) method, controlling for county and year of the second trimester. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by county-year-of-second-trimester cells, and cells are weighted by the number of births in each cell, or in the case of the birth rate, the number of women aged 15-44. Not all states report education or marital status, so these samples are smaller. Standard errors are clustered at the county level. $*\ p<0.10$, $**\ p<0.05$, $***\ p<0.01$.

Table 3: Robustness to Specification Changes: The Effect of WIC on Low Birth Weight

	Sun and Abraham (2021)						TWFE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A: Black Mothers							
WIC-month year-trim2	-0.3078***	-0.3038***	-0.3076***	-0.2787***	-0.2248**	-0.2515***	-0.3005***
	(0.1095)	(0.0898)	(0.0862)	(0.0892)	(0.0941)	(0.0872)	(0.1054)
Control Y-mean	13	13	13	13	13	13	13
Adj. R-Squared	0.028	0.028	0.028	0.028	0.028	0.030	0.028
Obs	189417	189417	189417	189381	189381	189381	189417
B: White Mothers							
WIC-month year-trim2	-0.0585	-0.0337	-0.0314	-0.0269	-0.0130	-0.0208	-0.0512
	(0.0415)	(0.0234)	(0.0233)	(0.0239)	(0.0248)	(0.0242)	(0.0414)
Control Y-mean	6	6	6	6	6	6	6
Adj. R-Squared	0.053	0.054	0.054	0.054	0.054	0.055	0.053
Obs	440037	440037	440037	439989	439989	439989	440037
C,Year FE	Y	Y	Y	Y	Y	Y	Y
1970 controls-pscore weighting	Y						Y
FS+CHC exp			Y	Y	Y	Y	
Cnty Time Varying Conts				Y	Y	Y	
1970 controls*Year FE					Y	Y	
State*Year FE						Y	

Notes: Table reports the estimated coefficients and standard errors using the Sun and Abraham (2021) method and TWFE, controlling for county and year of the second trimester. Panel A presents results for Black mothers, and Panel B for White mothers. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells. Columns 2–6 report estimates from separate regressions without propensity scores with additional controls in progression. Column 7 produces estimates using TWFE with propensity scores. Standard errors are clustered at the county level. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

Table 4: Magnitudes of Treatment Effects on the Treated, This Paper and Others

	Low birth weight			Birth weight		
	Reduced form on outcome	First stage on participation	TOT (col. 1/col. 2)	Reduced form on outcome	First stage on participation	TOT (col. 1/col. 2)
Panel A: *Infants of Black mothers*						
WIC, our paper	-0.3078	0.36	-0.855	11.7089	0.36	32.52
Food stamps, Almond et al. (2011)	-0.15	0.41	-0.366	3.454	0.41	8.424
Panel B: *Infants of White mothers*						
WIC, our paper	-0.0585	0.15	-0.39	5.1234	0.15	34.156
Food stamps, Almond et al. (2011)	-0.06	0.13	-0.462	2.039	0.15	13.593

Notes: Table reports the reduced form estimates for our paper and for Almond et al. (2011)) and first stage on participation (Figure 2), as well as the implied back-of-the-envelope treatment effects. Columns 1–3 are for the outcome low birth weight, and Columns 4–6 are for birth weight. Panel A is for infants of Black mothers and Panel B is for infants of White mothers.

Appendix

A Procedure for Digitizing Data from the National Archives

- Scanned and took pictures of plans at the National Archives in D.C. Members of the research team went to the National Archives, in Hyattsville MD, and requested archival materials that included State Plans for WIC (required after 1977 for some states, and 1978 for all states and Indian Tribal Organizations operating WIC). They also contained various memos from USDA to States and the USDA regional offices. We photographed all the information about WIC locations, discussed below, and then digitized it, as described below. Note that the state plans allowed us to fill in many counties/states excluded from Hoynes et al. (2011)'s compilation for years between 1978 and 1981 for some states not included in the HPS data and for the year 1981. .

- FIPS code changes: Counties can change their boundaries, with new counties being created by merging existing counties, splitting existing counties, and combining independent cities with their surrounding counties. We exclude Hawaii, Alaska and Virginia from our analysis. We constructed "Super Counties" for our control measures by these splitting and merging counties, where the new "Super County" is the smallest consistent county, merged onto 2016 FIPS codes and county definitions. So, for example, the Yuma-La Paz "Super County" includes Yuma county until 1983, after which Yuma county was split into two counties, La Paz and Yuma. For 1983 and after, Yuma and La Paz are combined into one "Super County". The BEA, REIS and SEER used slightly different county codes across part of the time period, we have mapped these to standard Census FIPS codes.

- In addition to the state plans, data sources used to digitize WIC implementation included the following:
  - To Save the Children Nutritional Intervention Through Supplemental Feeding (FY 1974): Included the USDA announcement of first grantees on August 1973
  - Maternal, Fetal, and Infant Nutrition (FY 1974): Listed sites selected to participate in the Special Supplemental Food Program for Women, Infants, and Children, prepared in December 1973
  - Agriculture-Environmental and Consumer Protection Appropriations for 1975 (FY 1974): Listed the USDA approved WIC Project Area Summary of Grants as of March 1974
  - Women and Children First, or Last? A report on the Special Supplemental Food Program for Women, Infants and Children, Prepared by Virginia Fleming, The Children's Foundation, Washington DC (FY 1975): Listed WIC sites as of April 1975, plus an update of an additional 45 grantees
  - Special Supplemental Food Program for Women, Infants, and Children, WIC program directory of local agencies (FY 1978): Included lists of local agencies serving migrant populations

  - Special Supplemental Food Programs for Women Infants and Children, The Children's Foundation, Washington DC (FY 1979): Published a list of participating counties by state
  - State Plans (FY 1979–FY 1981): Each state published lists and/or maps that listed WIC agencies currently participating and counties served, counties planning to participate, and counties not participating; some states included information by county of special populations served. The specific information included and how it was presented varied by state.
  - Impact of administration's economic proposals on programs under the jurisdiction of the Education and Labor Committee (FY 1981): Hearings before the Committee on Education and Labor, House of Representatives, Ninety-seventh Congress, first session, hearings held in Washington, D.C. held between March and May 1981 that included lists of counties served by WIC on January and September 1980
- Undergraduate research assistants and some consultants hand-entered all available data from the National Archives and other sources into Excel spreadsheets including: Plan date, state, clinic address, county name, county FIPS, special populations served, and the number of counties served as available from the sources listed above (special populations included Native Americans or Migrants)
- The research team then converted the digitized data by state and source into STATA datasets with indicator variables for WIC participation by fiscal year and specific date (if available), special population(s) served, and variables to flag counties that required additional scrutiny to determine WIC participation
- The resulting datasets were validated in a number of ways
  - Counties participating in WIC per our digitization were compared to counties participating in WIC from earlier research generously shared with us by Hoynes et al. (2011), where the Migrant Directory and Children's Book were used to identify which counties participated in WIC; each discrepancy was investigated by the research team to generate a shared decision about the inclusion (or exclusion) of a given county from our data
  - Data was assessed for consistency across years with additional validation for counties that had and subsequently lost WIC
  - Data was validated across and within sources. For example, many state WIC plans included a list of participating counties and agencies in addition to map(s) that indicated participation; due to the risk for error with hand-entered data on the part of the original state agencies and our subsequent digitization, sometimes the counties included in the list were different from those indicated by a map. In all cases after confirming correct data entry, we looked for a third source of information to confirm the presence or absence of WIC when two sources disagreed; each county that had a discrepancy was flagged for inconsistency and detailed information about the issue was recorded in a log.

  - The research team met regularly to review the digitization process; in those meetings, we discussed each state's counties with WIC by fiscal year and assessed any counties with inconsistencies
  - Regarding inconsistent source reporting of WIC, the research team reviewed each original source of information and established rules for indicating the presence of WIC in a given county in a given fiscal year; all counties with differing information were flagged as "county had internal conflict within source" as follows:"
    * If at least two out of three sources of information indicated the presence of WIC, the county in question was coded as having WIC
    * If only a list and map that disagree are available in state WIC plans, we assumed that information from a list of WIC agencies/counties was more reliable than a map where shading and a map key may not be as accurate, see example below.
    * If only two sources of information are available and they disagree regarding the presence of WIC in a county, the county was coded as having WIC if WIC was present in the county in the prior fiscal year.
    * We also individually assessed counties where the source information was unclear (i.e., smudged and/or hard to read) and where information across years was inconsistent, e.g., a county was indicated to have WIC in fiscal year t but not in $t + a$. These counties were flagged as "problematic" if we were unable to determine the presence of WIC with certainty.
    * Each county was reviewed with all members of the research team evaluating the source information to determine the best likelihood of the presence or absence of WIC in a given county in a given fiscal year.
  - The digitization and review process resulted in 13 problematic county-years and 22 county-years with internal conflicts.
- Plans also included information on special populations that were served by WIC agencies. We created variables that indicated the presence of WIC for any individual in the county and a variable that excluded counties where the only population served was a special population. These included:
  - Migrant families
  - Military families
  - Native American families
- Some Native American Reservations had WIC clinics that served Native American residents; the program area for these programs were mostly run by Indian Tribal Organizations and thus were not always contained within county boundaries; a county that included a clinic that served families of a specific Indian Tribe or Tribes was determined to have WIC. These counties were also flagged to indicate that only a portion of the county area was served by WIC.

- A detailed explanation of the digitization process by fiscal year and state including raw source information for counties with conflicting WIC indications was maintained by the research team; this includes all internal conflict counties, special population counties, and decisions regarding the presence of WIC.

B Sample Restrictions

We begin by using all the observations in the Vital Statistics Natality Data. We exclude births to mothers who are Native American; we are exploring these birth outcomes in another paper that is focused on WIC as operated by Indian Tribal Organizations or otherwise experienced by infants born to Native American mothers. We include some results for births to non-Native American, non-Black, and Non-White mothers (Other mothers, including Asians) in the Appendix tables. The results for Other mothers tend to be bounded in magnitude between those for Blacks and Whites (Appendix Table A.8).

Earlier states and years of the data are a 50% sample of birth certificates, so we weight these observations by 2. We exclude infants of mothers who do not reside in a US state.

We code WIC exposure according to WIC availability in the county of residence at the time of birth during the year of the mother's second trimester of pregnancy.

When gestational age at birth is unavailable, we impute it to be 40 weeks, which includes a number of states and years of occurrence as well as births where this was not reported for a specific birth certificate but the state of occurrence reported it.

Marital status, education, and some components of the measure of adequacy of prenatal care are not available for all states and years of occurrence.

C Infant mortality outcomes

This appendix contains results from running our preferred specification using alternate outcome variables about infant death. Infant mortality can be a more complex result to interpret than other outcomes. First, the predicted sign could go in either direction depending on the mechanism of action. Based on the results in the main body of the paper, we see improvements in birth weight outcomes for infants exposed to WIC in-utero. This could lead to a population of babies born at more robust birth weights that are more likely to survive an adverse event in the beginning of life. Alternatively, this could lead to more live births where the fetus otherwise would not have survived in-utero without the presence of WIC. These infants could be too weak to survive much longer after birth, leading to an increase in neonatal mortality as a result of WIC exposure. Second, the presence of WIC is still impacting the infants after birth. Families are eligible for WIC both during pregnancy and while children are young, so the infants who benefited from WIC in-utero may still be benefiting from WIC during their first year of life. Separating out deaths in the first month of life from deaths in the first year of life can give further insight into results that are a by-product of in-utero exposure versus further exposure during the first year/things that happened during birth or after. Because deaths from malnutrition in the first year of life are rare in the United States in this time period, we would not expect to see much infant mortality occuring during this time period.

Data from the Vital Statistics birth records described previously are used along with data from the Vital Statistics death records from 1968 to 1981. The data are similar to Vital Statistics birth records but are created from individual death records. They represent every single death in each state for every year except 1972, which is a 50% sample of deaths in each state. These data are used to construct the outcome variables for infant mortality (death during the first year of life), neonatal mortality (death during the first month of life), and post-neonatal mortality (death between months two to twelve). Rates are calculated by dividing the number of deaths by the number of births in a cell of the year of the start of the second trimester, month of the start of the second trimester, reported race of the infant, and county. Results are focused on Black and White infant deaths.

The methods follow the preferred specification in the main body of the paper but are robust to a variety of different specifications. Propensity-score weights at the cell level are calculated that predict WIC participation based on pre-WIC county characteristics and other pre-WIC program participation. Cells are adjusted using these weights as described in the main text. The results presented in this appendix follow the Sun and Abraham (2021) approach in order to account for potential bias in a staggered treatment adoption context with heterogeneity in treatment effects. The variation in the rollout of WIC over time and across counties is used with a difference-in-differences specification to estimate the causal effect of WIC access during pregnancy on infant mortality. Following Sun and Abraham (2021), the control group is counties that were "never-treated" (during the time span of our analysis).

The equation is as follows:

$$Y_{rct} = \alpha + \beta \cdot WIC_{ct} + \gamma_c + \tau_t + \epsilon_{rct} \tag{2}$$

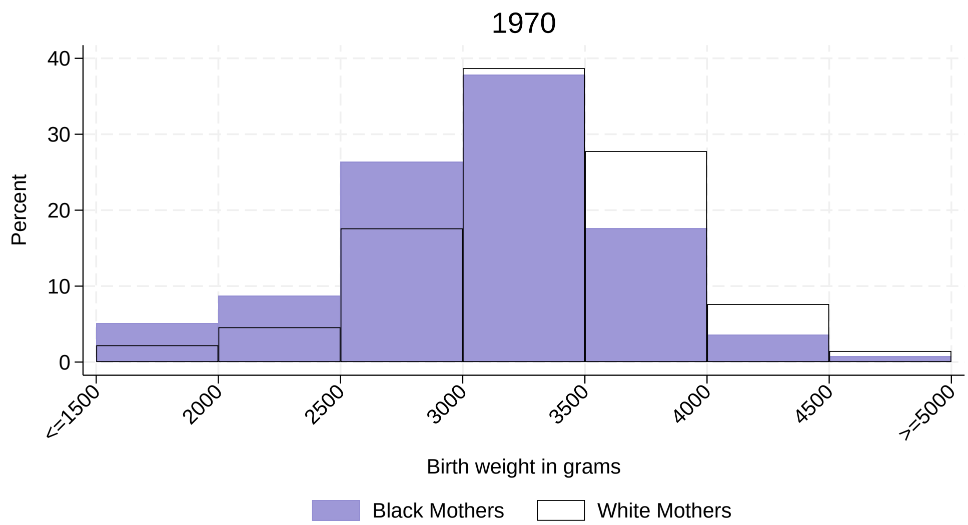
where Y_{rct} is the mortality outcome by cell for race r in county c during the month and year of the beginning of the second trimester t. Mortality outcomes are infant mortality, neonatal mortality, and post-neonatal mortality rates, where the number of neonatal deaths and post-neonatal deaths equal the number of infant deaths by definition. WIC_{ct} is equal to 1 if WIC is available in that county at that time. γ_c are county fixed effects and τ_t are month and year of the second trimester fixed effects. Cells are weighted by the propensity-score weights described previously. The regression is run separately for Black and White infant deaths.

As in the results in the main text, one of the contributions of this work is having rich enough data to separate out effects by race, where we see different patterns in outcomes as a result of the introduction of WIC as discussed in the main text. Figures A.13 (a) through (c) show infant mortality (a), neonatal mortality (b), and post-neonatal mortality (c) for Black infants. The intent-to-treat effect of WIC on infant mortality rates for Black infants is flat, but when overall infant mortality is decomposed into neonatal mortality (first month) and post-neonatal mortality (after the first month but within the first year), it becomes apparent that there is a small, but not statistically significant, increase in the neonatal mortality rate that is canceled out by a decrease in the post- neonatal mortality rate that causes overall effects on infant mortality rates to look flat for blacks. Figures A.14 (a) through (c) show the same set of results for White infants, where we see a different effect of WIC on the White infant mortality rate. Here, there is an increase in the neonatal mortality rate (b) and the post-neonatal mortality rate (c) remains flat, leading to an overall increase in the infant mortality rate (a). As we saw in the main set of results, we see a positive result for Black infants, the decrease in the post-neonatal mortality rate, that does not show up for the White population. As Black populations are more likely to be disadvantaged, there may more space for improvements to post- neonatal mortality rates as a result of WIC. The White neonatal mortality rate (b) shows some evidence of pre- trends during this time period. These specific results may be considered with some caution, but, overall, the set of all the results taken together paint a broad picture consistent with the main results.

Overall, we see results that tell a coherent story. The introduction of WIC at the start of the second trimester leads to a small increase in neonatal mortality, which is likely a consequence of a positive effect of WIC in-utero that leads to more births in place of stillbirths. This hypothesis is strengthened by the finding that post-neonatal mortality rates are either flat or decreasing.

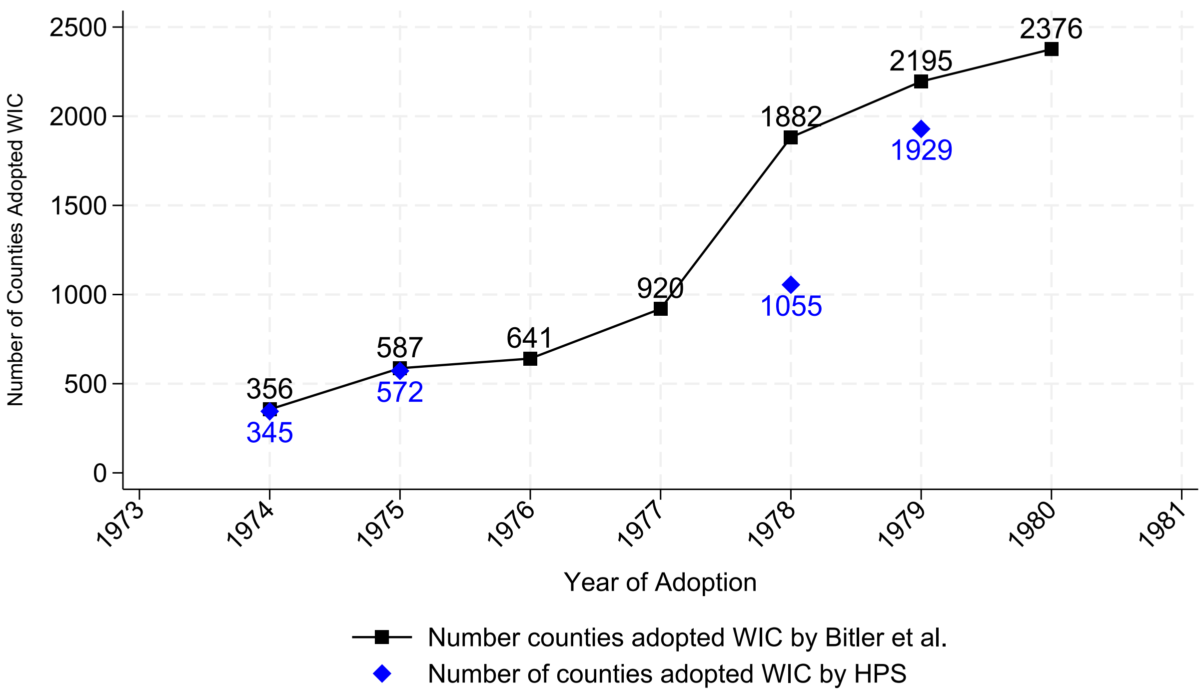
D Appendix Figures

Figure A.1: Birth Weight Distribution by Race, 1970



Notes: The data are from the Vital Statistics Natality Files. The figure reports the percentage of births within various bins by race.

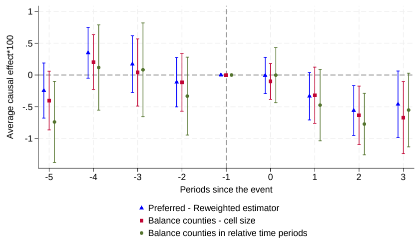
Figure A.2: Number of Counties Adopting WIC over Time, between 1974 and 1980: Our Data Compilation vs. Hoynes et al. (2011)



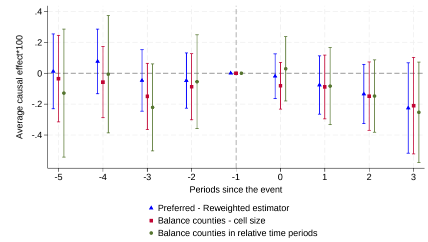
Notes: Hoynes et al. (2011) (HPS) generously shared their data for the counties included in their analysis. HPS had information for 1978 for only 13 states' with migrant programs and had information for 1979 only from a WIC directory. The State plans we have added include detailed information on local agencies and on all the counties served by WIC agencies and we have added WIC rollout information for the years 1978, 1979, and 1980.

Figure A.3: Robustness Checks for the Event Study Analysis: The Effect of WIC on Main Birth Outcomes—Black Mothers

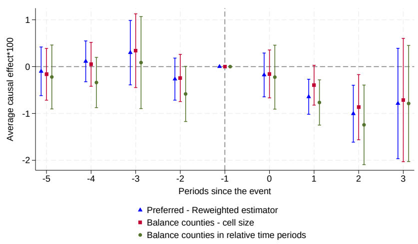
(a) Low Birth Weight



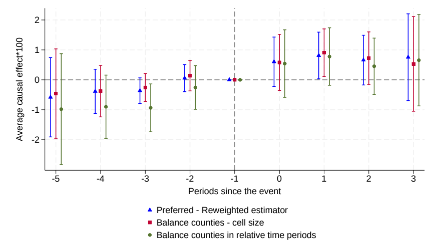
(b) Very Low Birth Weight



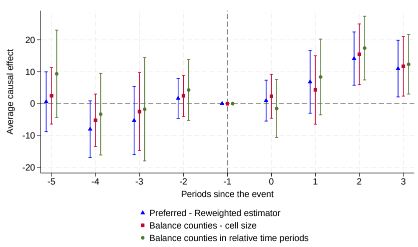
(c) Small for Gestational Age



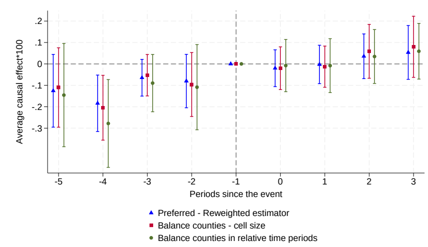
(d) Preterm



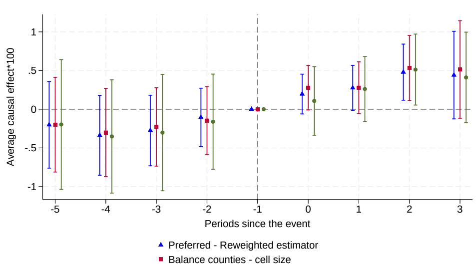
(e) Birth Weight



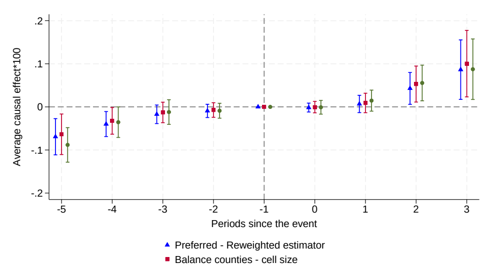
(f) High Birth Weight



(g) Large for Gestational Age



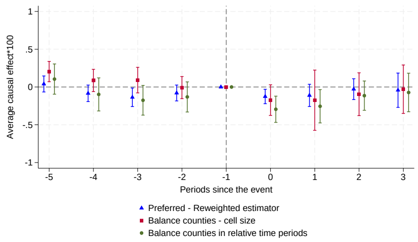
(h) Kessner Index Adequate



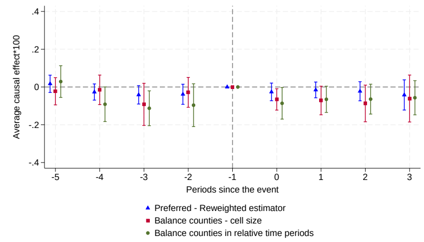
Notes: Figures report the estimated coefficients and their 95% confidence intervals using the Sun and Abraham (2021) method. All the dependent variables are multiplied by 100 (except for birth weight, which is measured in grams). Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. In blue (left most), we show our preferred specification. In red (middle), we present estimates using a balanced set of counties that have at least 2 Black and White births each month and year of the second trimester. This sample represents 85% of the total Black births in our preferred specification. In green (right most), we show estimates based on a balanced set of counties across leads and lags. This sample represents 89% of the total Black births in our preferred specification. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level. We binned relative time periods more than 6 years before and more than 4 years after WIC adoption. Standard errors are clustered at the county level.

Figure A.4: Robustness Checks for the Event Study Analysis: The Effect of WIC on Main Birth Outcomes—White Mothers

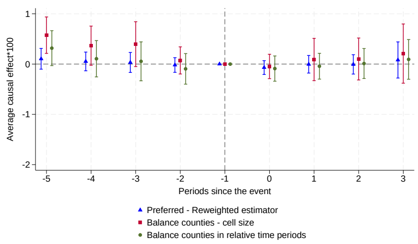
(a) Low Birth Weight



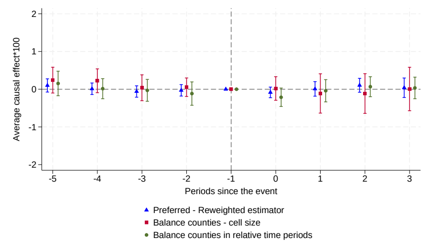
(b) Very Low Birth Weight



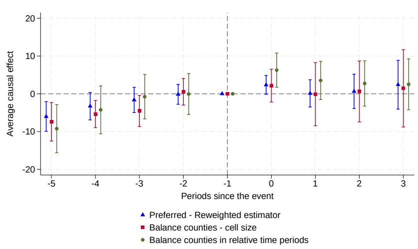
(c) Small for Gestational Age



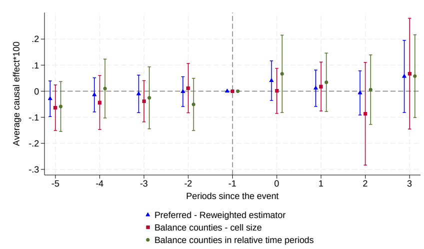
(d) Preterm



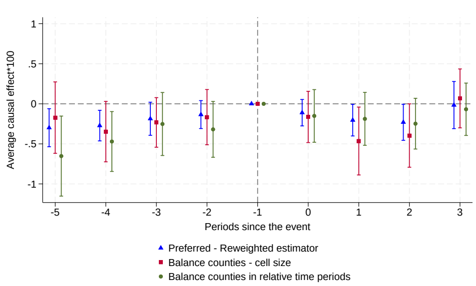
(e) Birth Weight



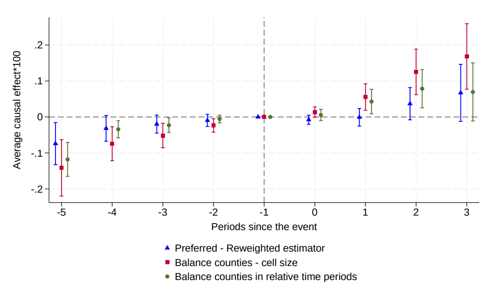
(f) High Birth Weight



(g) Large for Gestational Age

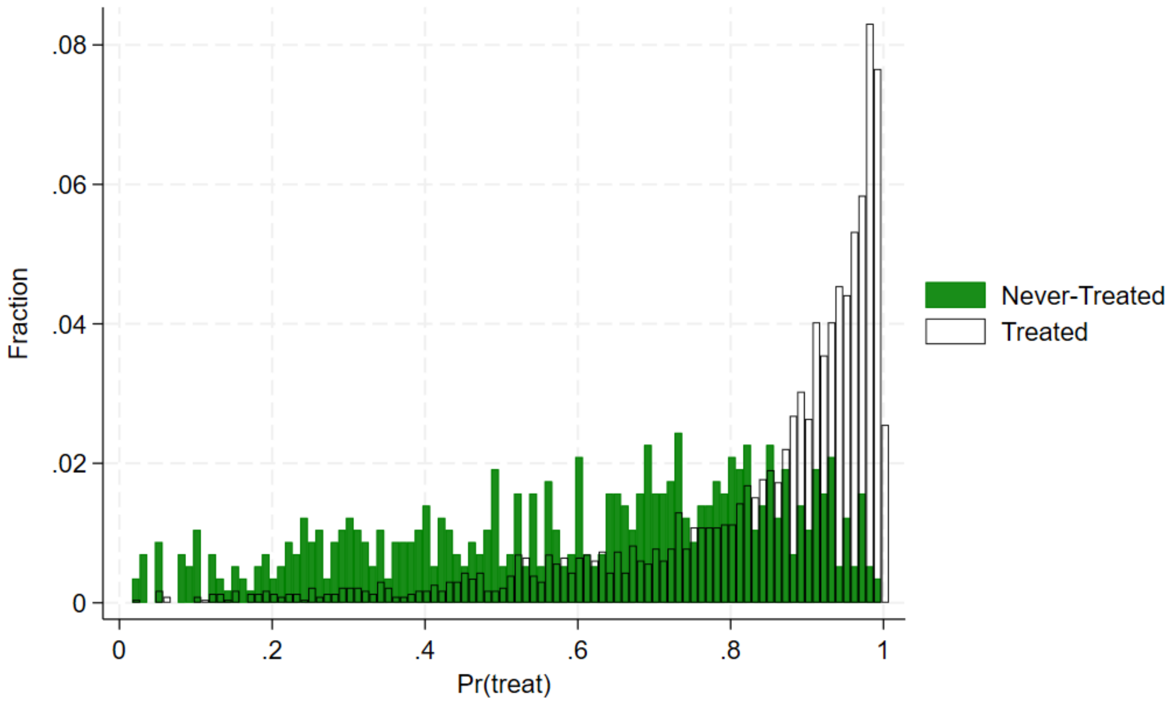


(h) Kessner Index Adequate



Notes: Figures report the estimated coefficients and their 95% confidence intervals using the Sun and Abraham (2021) method. All the dependent variables are multiplied by 100 (except for birth weight, which is measured in grams). Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. In blue (left most), we show our preferred specification. In red (middle), we present estimates using a balanced set of counties that have at least 2 Black and White births each month and year of the second trimester. This sample represents 56% of the total White births in our preferred specification. In green (right most), we show estimates based on a balanced set of counties across leads and lags. This sample represents 87% of the total White births in our preferred specification. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level. We binned relative time periods more than 6 years before and more than 4 years after WIC adoption. Standard errors are clustered at the county level.

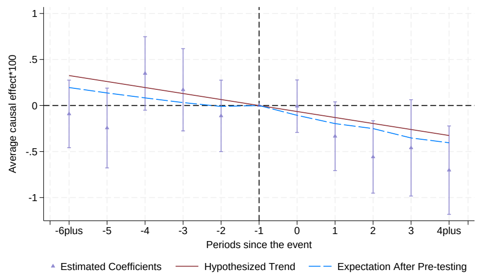
Figure A.5: Common Support: Histogram of Propensity Scores for Treated and Never-Treated Counties (during Sample Period)



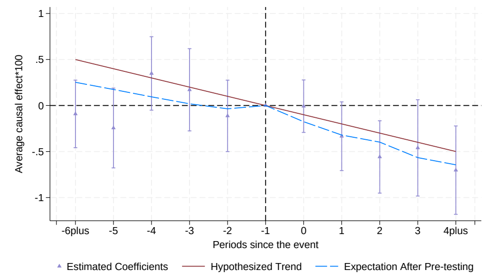
Notes: Figure reports the distribution of the predicted propensity to adopt WIC using pre-program county characteristics as predictors. To calculate the predicted probability, we run a logistic regression of WIC adoption on county pre-program characteristics, weighted by the number of births in the county measured in 1970. Pre-program characteristics include: Percent of the population below poverty, percent of the population at or above 65, percent of population with 4 years of schooling, percent with more than 12 years of schooling, percent non-White, percent urban, percent rural farm, Medicare and Military Health spending per capita, state UI spending per capita, the natural log of population, the number of medical doctors per capita, a dummy for Food Stamp adoption before 1974, a dummy for Community Health Center adoption before 1974, the state rate of White and non-White women and children in AFDC (averaged between 1967 and 1973), the infant mortality rate, the percent of births that are by teen mothers, the percent of births that are low birth weight, the number of hospitals per capita, a dummy for the Community Supplemental Food Program ever having been adopted, and the percent of first births,

Figure A.6: Event Study Roth (2022) Pre-Trends: The Effect of WIC on Main Birth Outcomes—Black Mothers

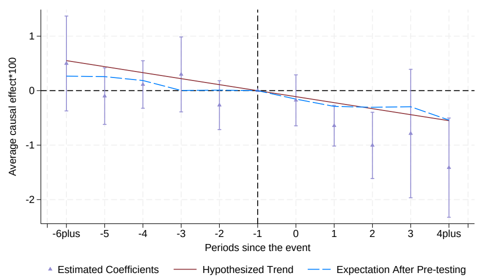
(a) Low Birth Weight—Power 0.5



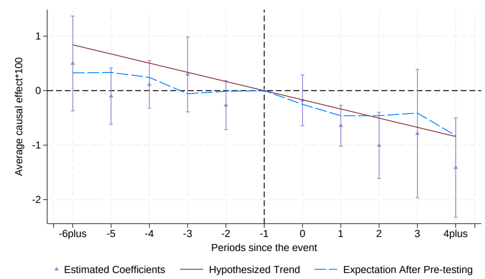
(b) Low Birth Weight—Power 0.8



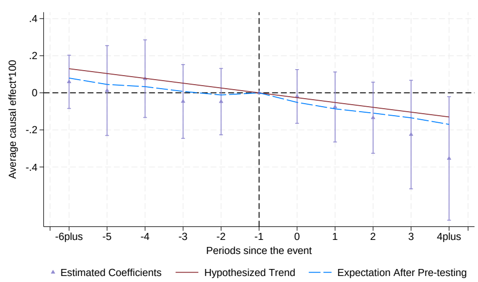
(c) Small for Gestational Age—Power 0.5



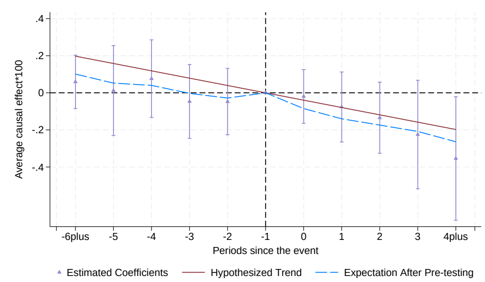
(d) Small for Gestational Age—Power 0.8



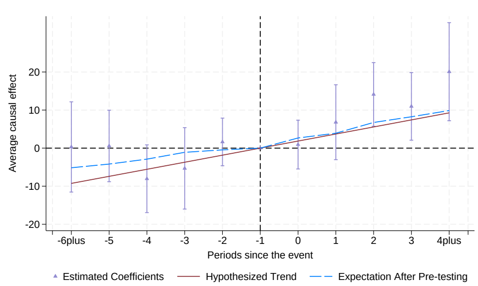
(e) Very Low Birth Weight—Power 0.5



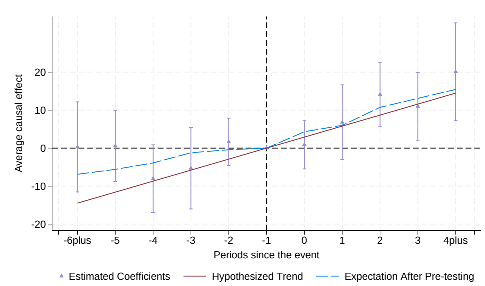
(f) Very Low Birth Weight—Power 0.8



(g) Birth Weight—Power 0.5



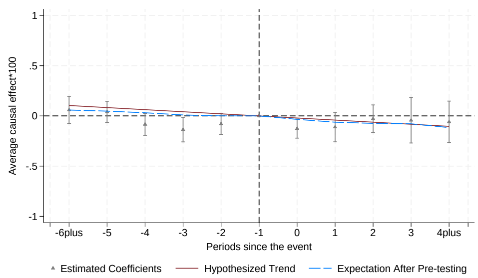
(h) Birth Weight—Power 0.8



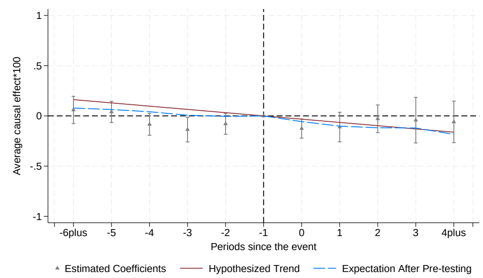
Notes: Figures report the estimated coefficients and their 95% confidence intervals using the Sun and Abraham (2021) method. In addition, we plot potential violations of parallel trends based on methods from Roth (2022) and report a p-value for a joint significance test for the pre-period coefficients being different from zero. The red solid line represents a hypothesized linear violation of the parallel trends with 50% and 80% power. The blue dashed line plots the expected pattern of the event studies coefficients on average, conditional on not detecting a significant pre-trend when in fact the true pre-trend was the hypothesized red line. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption based on pre-program characteristics. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level. Standard errors are clustered at the county level.

Figure A.7: Event Study Roth (2022) Pre-Trends: The Effect of WIC on Main Birth Outcomes—White Mothers

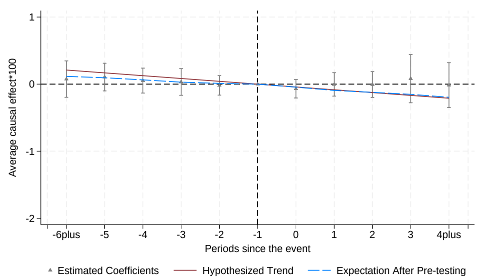
(a) Low Birth Weight—Power 0.5



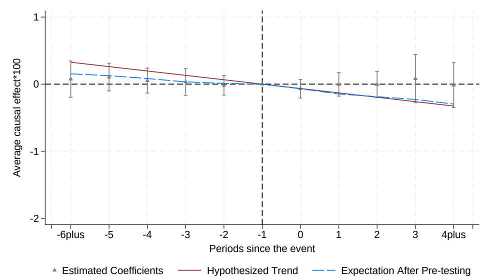
(b) Low Birth Weight—Power 0.8



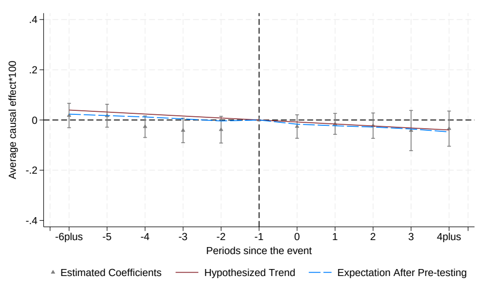
(c) Small for Gestational Age—Power 0.5



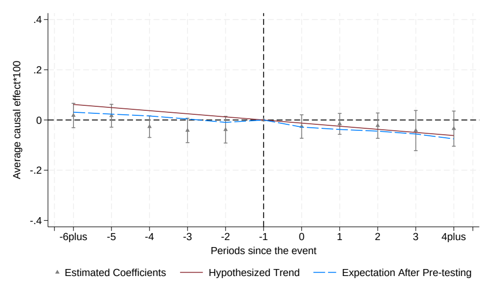
(d) Small for Gestational Age—Power 0.8



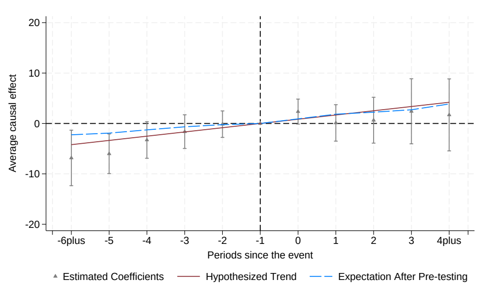
(e) Very Low Birth Weight—Power 0.5



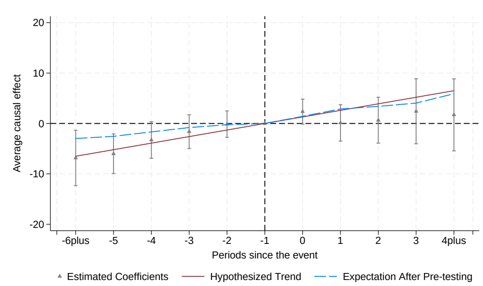
(f) Very Low Birth Weight—Power 0.8



(g) Birth Weight—Power 0.5

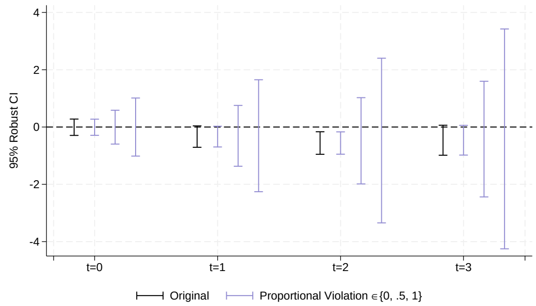


(h) Birth Weight—Power 0.8

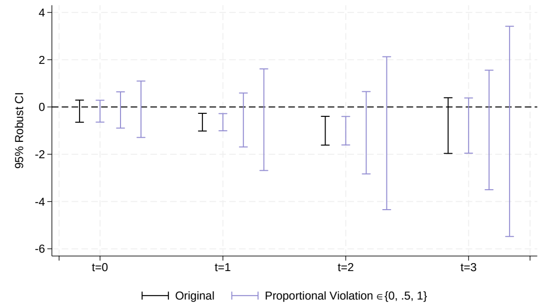


Notes: Figures report the estimated coefficients and their 95% confidence intervals using the Sun and Abraham (2021) method. In addition, we plot potential violations of parallel trends based on methods from Roth (2022) and report a p-value for a joint significance test for the pre-period coefficients being different from zero. The red solid line represents a hypothesized linear violation of the parallel trends with 50% and 80% power. The blue dashed line plots the expected pattern of the event studies coefficients on average, conditional on not detecting a significant pre-trend when in fact the true pre-trend was the hypothesized red line. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption based on pre-program characteristics. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells. Standard errors are clustered at the county level.

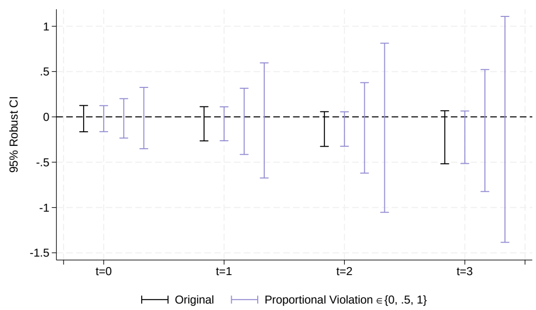
Figure A.8: Sensitivity Analysis using Rambachan and Roth, (2023): The Effect of WIC on Main Birth Outcomes—Black Mothers



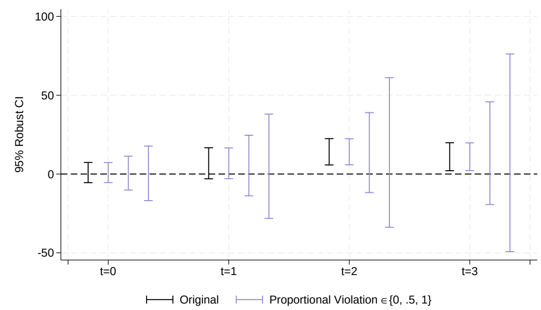
(a) Low Birth Weight



(b) Small for Gestational Age



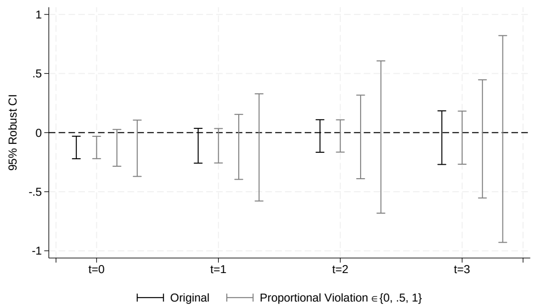
(c) Very Low Birth Weight



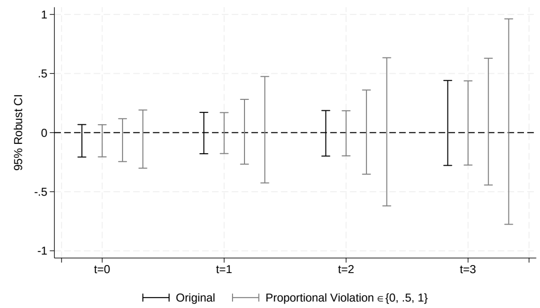
(d) Birth Weight

Notes: Figures present the results of Rambachan and Roth's (2023) approach to test the sensitivity of the event-study results. In purple, they show confidence intervals for the treatment effect at different event times using different multiples of possible post-treatment violations of parallel trends. In black, we plot our original results. A violation value equal to zero corresponds to allowing only linear violations of parallel trends, and larger values allow for larger deviations from linearity. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells.

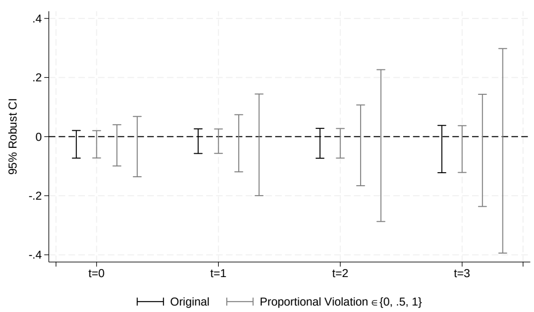
Figure A.9: Sensitivity Analysis using Rambachan and Roth, (2023): The Effect of WIC on Main Birth Outcomes—White Mothers



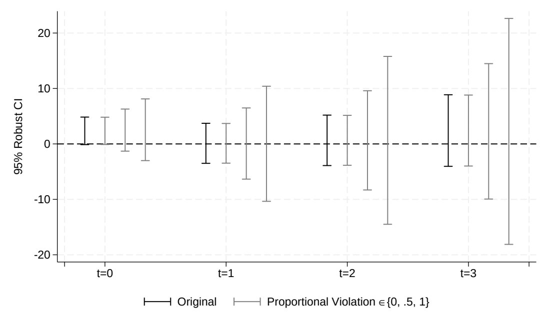
(a) Low Birth Weight



(b) Small for Gestational Age



(c) Very Low Birth Weight

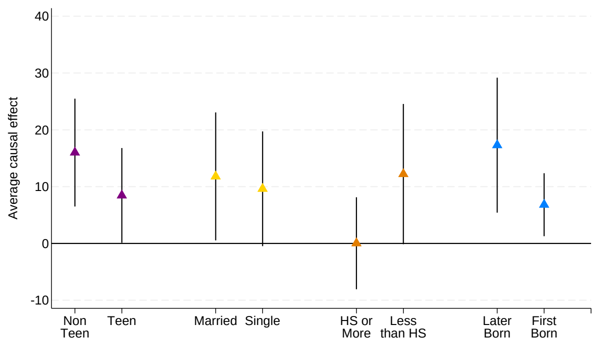


(d) Birth Weight

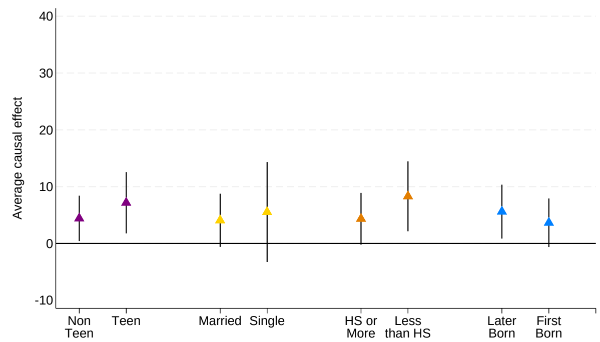
Notes: Figures present the results of Rambachan and Roth's (2023) approach to test the sensitivity of the event-study results. In purple, they show confidence intervals for the treatment effect at different event times using different multiples of possible post-treatment violations of parallel trends. In black, we plot our original results. A violation value equal to zero corresponds with allowing only linear violations of parallel trends, and larger values allow for larger deviations from linearity. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells.

Figure A.10: Heterogeneity Analyses: Individual Risk Factors—Birth Weight

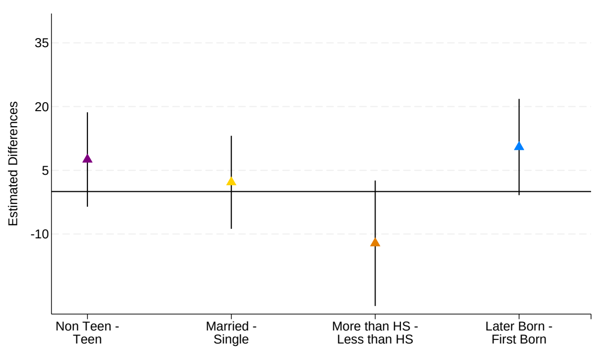
(a) Birth Weight—Black



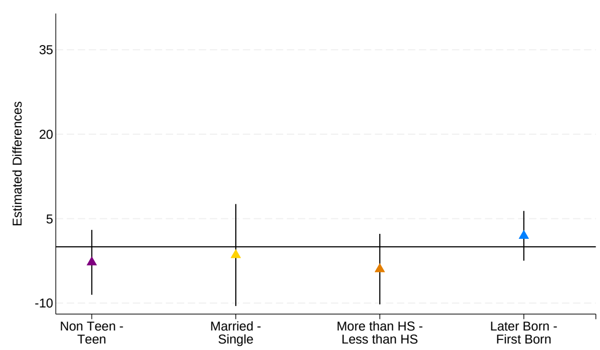
(b) Birth Weight—White



(c) Differences (more-less advantaged)—Black



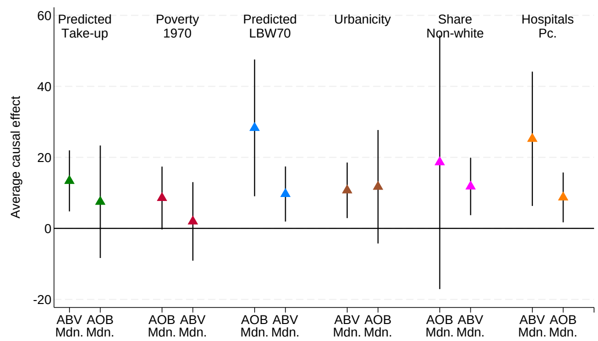
(d) Differences (more-less advantaged)—White



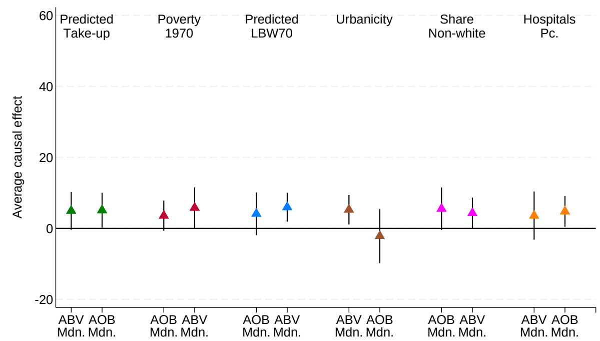
Notes: This analysis reports the estimated coefficients and the 95% confidence intervals, separately for each subgroup using the Sun and Abraham (2021) method. Each color represents a risk factor and corresponds to a separate regression. Each risk factor is defined based on the mothers' characteristics reported in the Vital Statistics. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. Panels A and B present estimates and 95% confidence intervals from fully saturated regressions, where each risk factor is fully interacted with all of the fixed effects within each race group. Panels C and D present the corresponding estimated difference between each factor group and its confidence interval. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by county, race, month, and year of the second trimester, parity, age category, gender, marital status, and education (finer cells than in our main analysis sample). Standard errors are clustered at the county level.

Figure A.11: Heterogeneity Analyses—County Risk Factors—Birth Weight

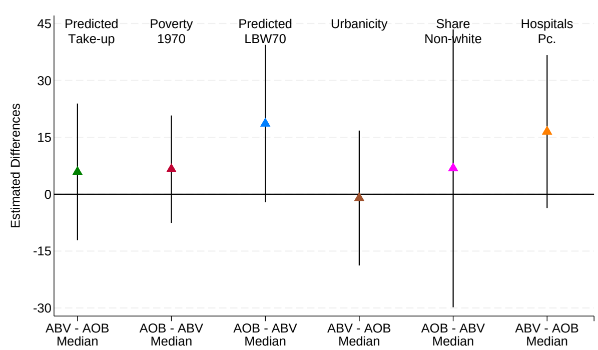
(a) Birth Weight—County Risk—Black



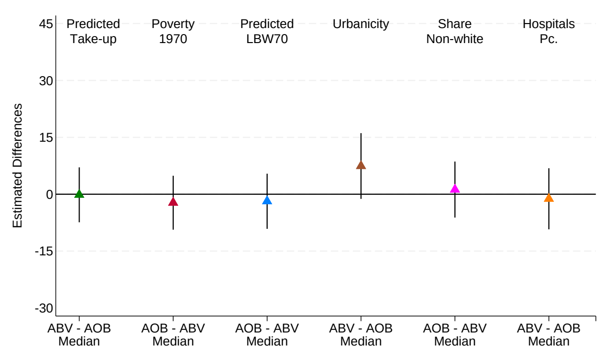
(b) Birth Weight—County Risk—White



(c) Differences across County Risk (more-less advantaged)—Black



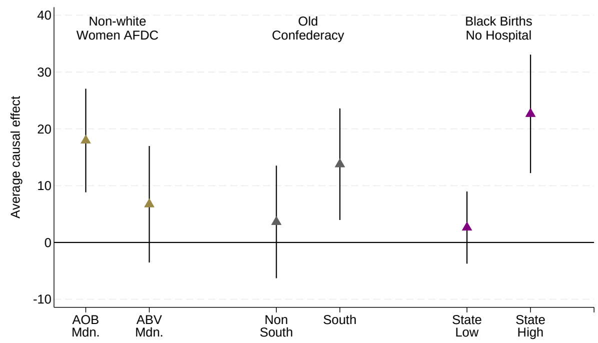
(d) Differences across County Risk (more-less advantaged)—White



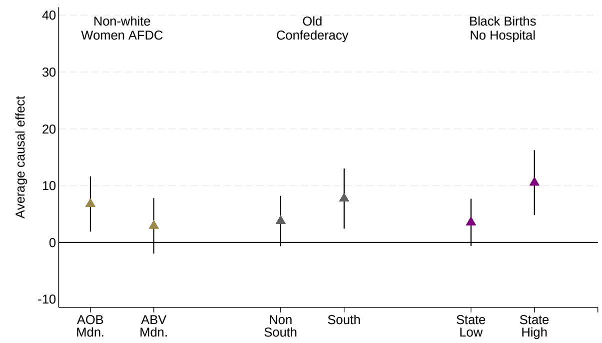
Notes: This analysis reports the estimated coefficients and the 95% confidence intervals, separately for each subgroup using the Sun and Abraham (2021) method. Each color represents a risk factor and corresponds to a separate regression. Each risk factor is defined based on county characteristics compiled using various data sources, as above median or at or below median county characteristics, with the unweighted median calculated using all counties after dropping AK and HI. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. Panels A and B present estimates and 95% confidence intervals from fully saturated regressions, where each risk factor is fully interacted with all of the fixed effects within each race group. Panels C and D present the corresponding estimated difference between each factor group and its confidence interval (i.e., counties at or below (AOB) the median of 1970 poverty rate and counties above the median (ABV)). We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by county, race, month and year of the second trimester, parity, age category, gender, marital status and education (finer cells than in our main analysis sample). Standard errors are clustered at the county level.

Figure A.12: Heterogeneity Analyses—State Risk Factors—Birth Weight

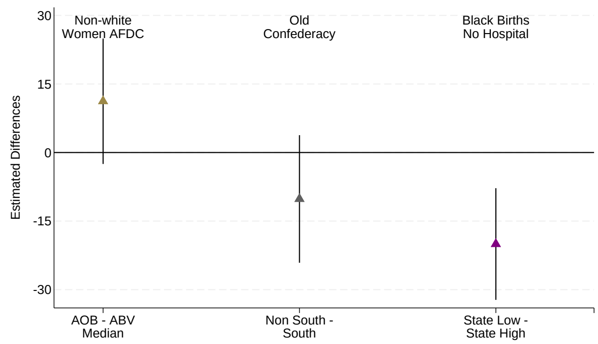
(a) Birth Weight—State Risk—Black



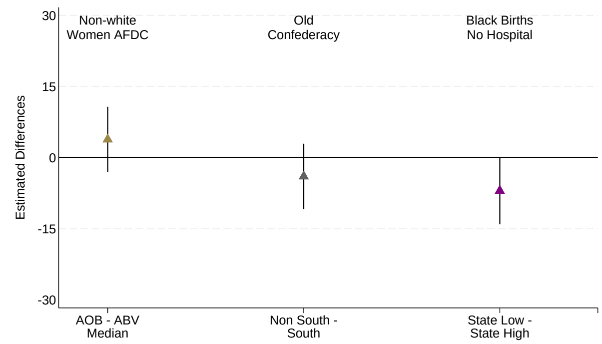
(b) Birth Weight—State Risk—White



(c) Differences across State Risk (more-less advantaged)—Black

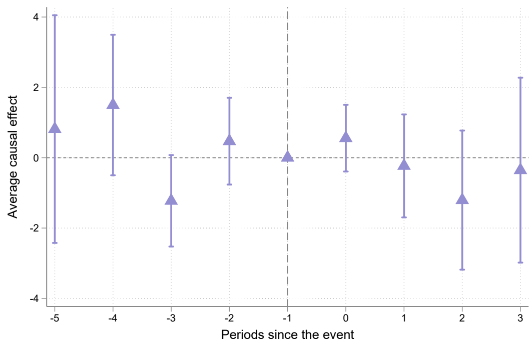


(d) Differences across State Risk (more-less advantaged)—White

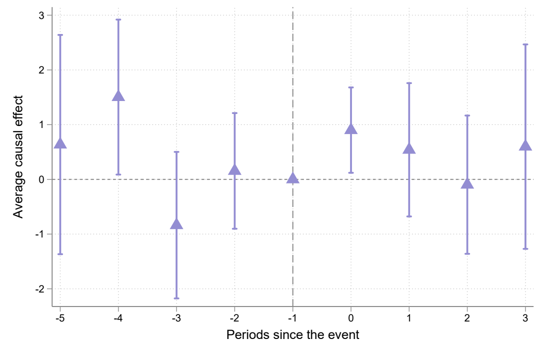


Notes: This analysis reports the estimated coefficients and the 95% confidence intervals, separately for each subgroup using the Sun and Abraham (2021) method. Each color represents a risk factor and corresponds to a separate regression. Each risk factor is defined based on state characteristics. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. Panels A and B present estimates and 95% confidence intervals from fully saturated regressions, where each risk factor is fully interacted with all of the fixed effects within each race group. Panels C and D present the corresponding estimated difference between each factor group and its confidence interval (i.e., states not of the old Confederacy and states of the old Confederacy). Non-white women in AFDC counts are from Goodman-Bacon (2018). States of the former confederacy include: Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia. The group of states with black births not in the hospital is based on the information reported in the Vital Statistics in 1968, as reporting above 5% of the births occurring not in the hospital. These states include: Alabama, Arkansas, Florida, Georgia, Mississippi, North Carolina, South Carolina, Texas, and Virginia. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by county, race, month and year of the second trimester, parity, age category, gender, marital status and education (finer cells than in our main analysis sample). Standard errors are clustered at the county level.

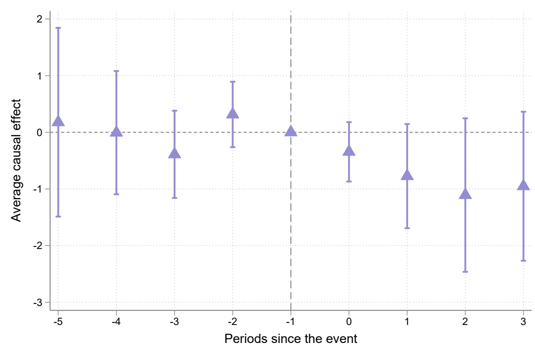
Figure A.13: Event Study: The Effect of WIC on Infant Mortality Outcomes—Black Mothers



(a) Infant Mortality



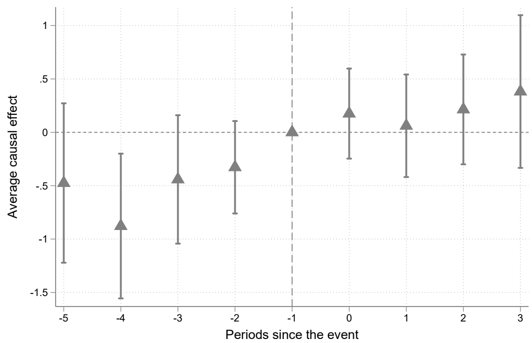
(b) Neonatal Mortality



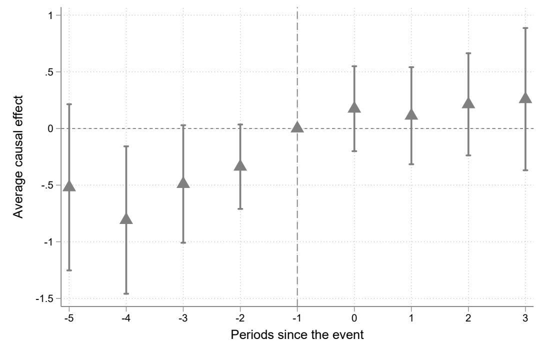
(c) Post-Neonatal Mortality

Notes: Figures report the estimated coefficients and standard errors using the Sun and Abraham (2021) method. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption based on pre-program characteristics. We use data from the Vital Statistics Natality Files that include births between 1968 and Q2, 1981 and Vital Statistics Mortality Files that include deaths between 1968 and 1980. We compile data on WIC timing from multiple sources between 1974 and 1981. The analysis is run using data at the cell level. We binned relative time periods more than 6 years before and more than 4 years after WIC adoption. Standard errors are clustered at the county level. $*\ p<0.10$, $**\ p<0.05$, $***\ p<0.01$.

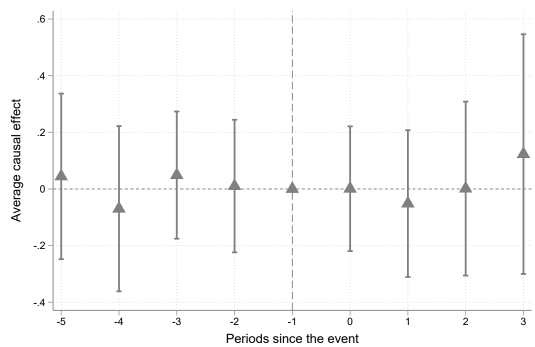
Figure A.14: Event Study: The Effect of WIC on Infant Mortality Outcomes—White Mothers



(a) Infant Mortality



(b) Neonatal Mortality



(c) Post-Neonatal Mortality

Notes: Figures report the estimated coefficients and standard errors using the Sun and Abraham (2021) method. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption based on pre-program characteristics. We use data from the Vital Statistics Natality Files that include births between 1968 and Q2, 1981 and Vital Statistics Mortality Files that include deaths between 1968 and 1980. We compile data on WIC timing from multiple sources between 1974 and 1981. The analysis is run using data at the cell level. We binned relative time periods more than 6 years before and more than 4 years after WIC adoption. Standard errors are clustered at the county level. $*\ p<0.10$, $**\ p<0.05$, $***\ p<0.01$.

E Appendix Tables

Table A.1: Summary Statistics for Individual Risk Factors, by Race

	Black	White
Individual Risk Factors		
Teen Mother	0.310	0.152
	(0.462)	(0.359)
Non-teen Mother	0.690	0.848
	(0.462)	(0.359)
Single Mother	0.483	0.071
	(0.500)	(0.257)
Married Mother	0.517	0.929
	(0.500)	(0.257)
Mother with Less than High School Degree	0.440	0.239
	(0.496)	(0.426)
Mother with High School Degree or More	0.560	0.761
	(0.496)	(0.426)
First Born Baby	0.390	0.417
	(0.488)	(0.493)
Later Born Baby	0.610	0.583
	(0.488)	(0.493)
N	1812940	7218881

Notes: We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. Averages for marital status and education variables only include the states that reported that information consistently during our study period. To calculate the averages, we use cells at the county, race, month and year of the second trimester, parity, age category, gender, marital status and education (finer cells than in our main analysis sample). N is the maximum number of cells used to calculate the averages. The sample excludes AK, HI and VA.

Table A.2: Cross-Tabulation of WIC rollout: Our paper vs. Hoynes et al. (2011)

	HPS Compitation							
Bitler et al. Compilation	1974	1975	1976	1977	1978	1979	New counties added by Bitler et al.	Total
1974	320	3	.	.	20	4	9	356
1975	18	196	.	.	6	9	2	231
1976	0	1	.	.	8	38	7	54
1977	0	5	.	.	85	151	38	279
1978	3	17	.	.	357	477	108	962
1979	0	2	.	.	3	161	147	313
1980	2	0	.	.	3	2	174	181
Missing in Bitler et al.	2	3	.	.	1	32	0	38
Total	345	227	.	.	483	874	485	2414

Notes: This table shows the cross-tabulation of two data compilations of WIC rollout based on the number of counties that adopted WIC in a given year. We describe our data compilation process in detail in Data Appendix A, and Hoynes et al. (2011) (HPS)'s data compilation is described in their paper. We note that HPS rollout years exclude 1977 and 1976 given they lack primary sources for that period. The sample corresponds to WIC counties in HPS or in our rollout data, but excludes Alaska, Hawaii, and Virginia.

Table A.3: Summary Statistics for Outcomes and Controls for Full Sample and by WIC Adoption Status, by Race

	All Counties		Counties with WIC		Counties without WIC	
	Black	White	Black	White	Black	White
Outcomes						
Low Birth Weight	13.203	6.225	13.190	6.221	13.388	6.272
	(23.088)	(14.266)	(22.972)	(13.887)	(24.704)	(18.266)
Very Low Birth Weight	2.375	0.917	2.385	0.920	2.237	0.884
	(10.364)	(5.713)	(10.335)	(5.582)	(10.778)	(7.116)
Small for Gest. Age	14.940	8.423	14.937	8.414	14.978	8.528
	(24.972)	(16.978)	(25.007)	(16.567)	(24.478)	(21.442)
Preterm	16.865	7.675	16.825	7.684	17.434	7.569
	(26.932)	(16.745)	(26.958)	(16.402)	(26.558)	(20.532)
Birth Weight	3086.844	3352.286	3086.867	3352.145	3086.510	3354.002
	(434.440)	(343.736)	(432.673)	(334.829)	(459.217)	(437.932)
High Birth Weight	0.732	1.762	0.726	1.759	0.826	1.801
	(6.271)	(7.535)	(6.208)	(7.305)	(7.118)	(9.922)
Large for Gest. Age	5.838	10.322	5.834	10.317	5.903	10.388
	(17.115)	(18.335)	(17.116)	(17.859)	(17.097)	(23.458)
Kessner Index - Adequate	0.282	0.400	0.282	0.395	0.274	0.454
	(0.356)	(0.398)	(0.357)	(0.395)	(0.340)	(0.426)
WOP exposure						
Food Stamps Exposure	0.900	0.863	0.902	0.865	0.873	0.830
	(0.300)	(0.344)	(0.297)	(0.341)	(0.333)	(0.375)
CHC Exposure	0.479	0.302	0.484	0.319	0.406	0.090
	(0.500)	(0.459)	(0.500)	(0.466)	(0.491)	(0.287)
Expenditures, Hospitals and Population						
UI State Expenditures per capita	125.283	132.639	128.760	137.003	75.061	79.478
	(110.510)	(110.786)	(112.593)	(111.734)	(52.966)	(81.525)
Medicare Expenditures per capita	270.227	258.707	271.631	259.865	249.972	244.599
	(135.015)	(117.327)	(136.800)	(117.505)	(103.883)	(114.199)
Number of Hospitals per capita	0.029	0.032	0.028	0.031	0.040	0.052
	(0.021)	(0.031)	(0.019)	(0.027)	(0.034)	(0.058)
Population (000s)	1160.478	782.416	1216.140	836.626	356.486	121.945
	(1790.881)	(1539.900)	(1837.118)	(1588.774)	(305.184)	(190.477)
1970 Characteristics						
1970 Percent Rural	0.027	0.047	0.027	0.042	0.036	0.101
	(0.051)	(0.077)	(0.051)	(0.073)	(0.054)	(0.104)
1970 Percent Living Below Poverty	13.808	10.607	13.454	10.518	18.929	11.693
	(9.128)	(6.821)	(9.193)	(6.840)	(6.186)	(6.487)
N	1812940	7218881	1680145	6275662	132795	943219

Notes: The outcomes are from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. The WOP exposure to Food Stamps and Community Health Centers comes from Hoynes et al. (2011) and Kose et al. (2024). The other controls come from Almond et al. (2011). To calculate the averages, we use cells at the county, race, month and year of the second trimester, parity, age category, gender, marital status and education (finer cells than in our main analysis sample). N is the maximum number of cells used to calculate the averages. The sample excludes AK, HI and VA.

Table A.4: Determinants of County WIC Rollout Timing as a Function of 1970 or Earlier County Characteristics

	Unweighted		Weighted		Unweighted		Weighted	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% population age <5-1970	0.018	0.039	-0.145	-0.219	0.017	0.040	-0.158	-0.195
	(0.044)	(0.047)	(0.137)	(0.145)	(0.044)	(0.047)	(0.130)	(0.145)
% population aged 65+ -1970	0.082***	0.064***	0.033	-0.030	0.081***	0.064***	0.008	-0.038
	(0.020)	(0.021)	(0.054)	(0.060)	(0.020)	(0.021)	(0.051)	(0.059)
Log 1970 population-1970	-0.211***	-0.376***	-0.541***	-0.506***	-0.213***	-0.375***	-0.551***	-0.539***
	(0.052)	(0.057)	(0.112)	(0.125)	(0.052)	(0.057)	(0.125)	(0.136)
% below poverty level-1969	-0.016*	-0.034***	-0.048**	-0.038*	-0.016**	-0.034***	-0.039**	-0.030
	(0.008)	(0.009)	(0.019)	(0.020)	(0.008)	(0.009)	(0.019)	(0.019)
% persons 25+ w/ <4 yrs sch-1970	-0.046***	-0.009	-0.034	0.003	-0.046***	-0.009	-0.044	-0.017
	(0.015)	(0.016)	(0.031)	(0.031)	(0.015)	(0.016)	(0.030)	(0.030)
% persons 25+ w/ 12+ yrs sch-1970	-0.002	0.015**	0.010	0.032**	-0.002	0.015**	0.008	0.029**
	(0.006)	(0.007)	(0.012)	(0.015)	(0.006)	(0.007)	(0.012)	(0.015)
% nonwhite-1970	0.002	-0.001	0.018	-0.001	0.002	-0.001	0.012	-0.007
	(0.003)	(0.004)	(0.014)	(0.009)	(0.003)	(0.004)	(0.013)	(0.009)
% urban-1970	-0.000	0.001	-0.011**	-0.018***	-0.000	0.000	-0.013***	-0.018***
	(0.002)	(0.002)	(0.005)	(0.005)	(0.002)	(0.002)	(0.005)	(0.005)
% rural farm-1970	-0.222	0.139	-1.295*	-2.861**	-0.240	0.122	-2.212***	-2.969***
	(0.380)	(0.470)	(0.780)	(1.170)	(0.383)	(0.470)	(0.842)	(1.121)
Medicare benefits pc (1000s 2006)-1970	-0.002*	-0.002*	-0.000	0.001	-0.002*	-0.002*	0.001	0.002
	(0.001)	(0.001)	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.003)
State UI compensation pc (1000s 2006)-1970	-0.002***	-0.001	-0.001	0.001	-0.002***	-0.001	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Hospitals pc-1970	-1.341**	-0.301	-3.878**	-2.017	-1.301**	-0.315	-4.250***	-2.200
	(0.587)	(0.565)	(1.658)	(1.775)	(0.590)	(0.567)	(1.624)	(1.708)
Active MDs pc -1960	-0.176**	-0.219***	-0.158	-0.262**	-0.183**	-0.219***	-0.218*	-0.313**
	(0.073)	(0.069)	(0.139)	(0.128)	(0.072)	(0.069)	(0.127)	(0.125)
Infant mortality rate-1970	0.002	-0.000	0.009	0.010	0.002	-0.000	0.011	0.011
	(0.003)	(0.003)	(0.008)	(0.007)	(0.003)	(0.003)	(0.008)	(0.007)
Adopted FS before 1974	-0.112	0.054	0.073	0.182	-0.081	0.052	0.149	0.279
	(0.094)	(0.115)	(0.190)	(0.221)	(0.097)	(0.116)	(0.194)	(0.205)
Adopted CHC before 1974	-1.170***	-1.056***	-0.746**	-0.570**	-1.184***	-1.057***	-0.888***	-0.646***
	(0.216)	(0.222)	(0.292)	(0.266)	(0.215)	(0.222)	(0.264)	(0.248)
CSFP before 1974					0.184	0.032	0.912***	0.776***
					(0.125)	(0.138)	(0.247)	(0.266)
CDP ever					0.054	-0.135	-1.067	-0.964
					(0.149)	(0.147)	(0.829)	(0.609)
Adj. R-Squared	0.107	0.222	0.327	0.428	0.107	0.221	0.356	0.443
Obs	2312	2312	2312	2312	2312	2312	2312	2312
State FE	N	Y	N	Y	N	Y	N	Y

Notes: The data are at the county level and the dependent variable is the year (normalized to be one in 1974) that WIC was introduced in the county, for counties adopting WIC in our time period. The covariates come from the City and County Data Book for 1970 and 1960 and other sources. Weighted regressions weight counties by the number of births in 1970. Heteroskedasticity-robust standard errors in parenthesis. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

Table A.5: Balance Test by WIC Adoption Status across Counties: 1970 County Characteristics Summary Statistics

	Unadjusted			Weighted by $\widehat{p}$-score			Add state FE		
	Mean Diff (1)	SE (2)	MHT p-value (3)	Mean Diff (4)	SE (5)	MHT p-value (6)	Mean Diff (7)	SE (clustered) (8)	MHT p-value (9)
% population age <5-1970	0.73***	0.07	0.01	0.06	0.11	0.99	0.15	0.12	1.00
% population aged 65+ -1970	-0.70***	0.07	0.01	-0.11	0.08	0.91	-0.05	0.16	1.00
Log 1970 population-1970	0.45***	0.03	0.01	0.13*	0.08	0.81	0.02	0.09	1.00
% below poverty level-1969	0.37***	0.05	0.01	-0.01	0.23	1.00	0.22	0.17	1.00
% persons 25+ w/ <4 yrs sch-1970	0.54***	0.05	0.01	-0.38	0.42	0.98	0.02	0.19	1.00
% persons 25+ w/ 12+ ysch-1970	-0.30***	0.05	0.01	-0.01	0.13	1.00	-0.16	0.12	1.00
% nonwhite-1970	0.43***	0.04	0.01	0.22**	0.11	0.70	0.09	0.13	1.00
% urban-1970	0.21***	0.05	0.01	0.02	0.08	1.00	0.07	0.13	1.00
% rural farm-1970	-0.87***	0.09	0.01	-0.14	0.17	0.98	-0.12	0.13	1.00
Medicare pc(1000s 2006)-1970	-0.56***	0.06	0.01	-0.17**	0.08	0.66	-0.04	0.10	1.00
State UI pc (1000s 2006)-1970	0.55***	0.03	0.01	0.13	0.11	0.98	-0.01	0.09	1.00
Hospitals pc-1970	-1.35***	0.16	0.01	-0.28	0.22	0.98	-0.15	0.18	1.00
Active MDs pc -1960	0.09***	0.02	0.01	0.08*	0.04	0.73	0.04	0.06	1.00
% births not hospital -1970	0.23***	0.06	0.01	-0.80	0.64	0.98	-0.27	0.32	1.00
Women AFDC rate 67-73 - White	0.36***	0.03	0.01	0.22***	0.08	0.48	0.00		
Women AFDC rate 67-73 - Nonwhite	0.29***	0.05	0.01	0.13	0.14	0.98	-0.00		
Children AFDC rate 67-73-White	0.33***	0.03	0.01	0.20***	0.07	0.41	0.00		
ChildrenAFDC rate 67-73-Nonwhite	0.27***	0.05	0.01	0.15	0.11	0.91	-0.00		
FS opened pre-1974	-0.11**	0.05	0.09	-0.08	0.08	0.98	0.03	0.09	1.00
CHC opened pre-1974	0.06***	0.01	0.01	0.06***	0.02	0.32	0.02	0.04	1.00
CSFP pre-1974	0.14***	0.02	0.01	0.08	0.06	0.98	-0.10	0.10	1.00
CDP ever	-0.13***	0.05	0.05	0.10	0.16	0.99	0.29	0.19	1.00
% low bw -1970	0.25**	0.10	0.05	0.32	0.21	0.89	0.03	0.24	1.00
Infant mortality rate-1970	0.15	0.11	0.18	-0.25	0.20	0.98	-0.08	0.16	1.00
% teen motherhood -1970	0.21***	0.06	0.01	-0.01	0.19	1.00	0.06	0.16	1.00
% first born -1970	0.20**	0.09	0.09	-0.21	0.16	0.98	-0.39***	0.13	1.00

Notes: Sample corresponds to all counties in our analyses. All county characteristics are standardized using z-scores. Mean differences correspond to the difference in average characteristics between counties with and without WIC. The characteristics come from the City and County Data Book for 1970 and 1960 and other sources. Weighted regressions weight counties by the number of births in 1970. Multiple hypothesis testing (MHT) correction uses Romano and Wolf (2005) procedure. Heteroskedasticity-robust standard errors in parenthesis. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

Table A.6: Joint Significance of Pre-trends, with and without Multiple Hypothesis Correction

		Family Wise Error Rate	**False Discovery Rate**
Outcome	***p*-values from regressions (1)**	***p*-values by Holm (1979) (2)**	***q*-values by Benjamini et al. (2006) (3)**
Panel A: Black Mothers			
Low Birth Weight	0.031	0.188	0.075
Very Low Birth Weight	0.423	0.502	0.190
Small for Gestational Age	0.039	0.197	0.075
Preterm	0.100	0.399	0.087
Birth Weight	0.167	0.502	0.126
High Birth Weight	0.019	0.135	0.073
Large for Gestational Age	0.208	0.502	0.135
Kessner Index for Adequate Prenatal Care	0.007	0.059	0.063
Panel B: White Mothers			
Low Birth Weight	0.001	0.010	0.006
Very Low Birth Weight	0.012	0.073	0.023
Small for Gestational Age	0.836	1.000	0.477
Preterm	0.242	0.726	0.164
Birth Weight	0.015	0.074	0.023
High Birth Weight	0.956	1.000	0.477
Large for Gestational Age	0.088	0.351	0.076
Kessner Index for Adequate Prenatal Care	0.000	0.000	0.001

Notes: Column 1 of this table reports the p-values from a joint test of statistical significance for all pre-period coefficients from Figures 3 and 4. In Column 2, we report p-values for a joint significance test of all pre-program coefficients after correcting for multiple hypothesis testing (MHT) using a step-down procedure introduced by Holm (1979) that controls for family-wise error rate. In Column 3, we report p-values that adjust for MHT using the Benjamini et al. (2006) procedure that accounts for the false discovery rate based on sharpened q-values.

Table A.7: de Chaisemartin and D'Haultfœuille (2020): Regression of TWFE weights (W_{gt}) on Factors Associated with The Treatment Effect—Low Birth Weight

	Correlation	Coef.	S.E	T-stat
Month_year	-0.727	-17.969	0.502	-35.761
Black	-0.021	-0.003	0.001	-4.326
State UI compensation per capita	0.291	43.499	5.315	8.184
Medicare benefits per capita	-0.250	-38.679	3.663	-10.560
Hospitals per capita	0.072	0.002	0.0002	10.303

Notes: We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. The sample corresponds to all births pooled by race. This table reports results using a diagnostic test introduced by de Chaisemartin and D'Haultfœuille (2020).

Table A.8: The Effect of WIC on Main Birth Outcomes: Other Race Mothers

	Low Birth Weight	Very Low Birth Weight	Small for Gest. Age	Preterm	Birth Weight	High Birth Weight	Large for Gest. Age	Kessner Index Adequate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WIC-month year-trim2	-0.5080	0.0074	-0.6517	0.5950	21.3981**	-0.1094	0.7925	0.0440*
	(0.5708)	(0.1821)	(0.7386)	(0.7878)	(9.8733)	(0.1388)	(0.6809)	(0.0258)
Control Y-mean	7.663	1.128	11.006	9.269	3194.771	0.666	5.800	0.225
% Effect (Coef/Control Mean)	-6.629	0.659	-5.921	6.419	0.670	-16.412	13.664	19.536
Adj. R-Squared	0.020	0.012	0.016	0.030	0.050	0.022	0.029	0.512
Obs	70275	70275	53287	53287	70275	70275	53287	57602
C,Year FE	Y	Y	Y	Y	Y	Y	Y	Y
1970 controls-pscore weighting	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Table reports the estimated coefficients and standard errors using the Sun and Abraham (2021) method, controlling for county and year of the second trimester for other race mothers. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group.We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells. Standard errors are clustered at the county level. $*\ p<0.10$, $**\ p<0.05$, $***\ p<0.01$.

Table A.9: Robustness to Specification Changes: The Effect of WIC on Birth Weight

	Sun and Abraham (2021)						TWFE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A: Black Mothers							
WIC-month year-trim2	11.7089***	11.6103***	11.5882***	10.0923***	5.0837*	9.4503***	11.1116***
	(4.0157)	(2.4526)	(2.4967)	(2.3835)	(2.6862)	(2.1671)	(3.8206)
Control Y-mean	3080	3080	3080	3080	3080	3080	3080
Adj. R-Squared	0.079	0.076	0.076	0.076	0.079	0.083	0.079
Obs	189417	189417	189417	189381	189381	189381	189417
B: White Mothers							
WIC-month year-trim2	5.1234***	3.9154***	3.7579***	4.1300***	3.1011***	1.7862*	4.6740**
	(1.8842)	(1.1007)	(1.1215)	(1.1091)	(1.1070)	(0.9306)	(1.8331)
Control Y-mean	3339	3339	3339	3339	3339	3339	3339
Adj. R-Squared	0.269	0.276	0.276	0.276	0.277	0.281	0.269
Obs	440037	440037	440037	439989	439989	439989	440037
C,Year FE	Y	Y	Y	Y	Y	Y	Y
1970 controls-pscore weighting	Y						Y
FS+CHC exp			Y	Y	Y	Y	
Cnty Time Varying Conts				Y	Y	Y	
1970 controls*Year FE					Y	Y	
State*Year FE						Y	

Notes: Table reports the estimated coefficients and standard errors using the Sun and Abraham (2021) method and TWFE, controlling for county and year of the second trimester. Panel A presents results for Black mothers, and Panel B for White mothers. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells. Columns 2–6 report estimates from separate regressions without propensity scores with additional controls in progression. Column 7 produces estimates using TWFE with propensity scores. Standard errors are clustered at the county level. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

Table A.10: Robustness to p-score Weighting: The Effect of WIC on Low Birth Weight and Birth Weight for All Counties and for Balanced Sample of Counties

	Low Birth Weight				Birth Weight			
	All Counties- p-score adj. · births	All Counties- p-score adj.	Balance Counties- p-score adj. · births	Balanced Counties- p-score adj.	All Counties- p-score adj. · births	All Counties- p-score adj.	Balance Counties- p-score adj. · births	Balanced Counties- p-score adj.
A: Black Mothers								
WIC-month year-trim2	-0.3078***	-0.0686	-0.2919**	-0.4870**	11.7089***	9.4562*	10.1074**	19.7008***
	(0.1095)	(0.2557)	(0.1151)	(0.2247)	(4.0157)	(5.3189)	(4.1205)	(6.0847)
Control Y-mean	13	13	14	14	3080	3080	3076	3076
% Effect (Coef/Control Mean)	-2.291	-0.510	-2.153	-3.592	0.380	0.307	0.329	0.640
Adj. R-Squared	0.028	0.016	0.056	0.026	0.079	0.040	0.143	0.077
Obs	189417	189417	71604	71604	189417	189417	71604	71604
B: White Mothers								
WIC-month year-trim2	-0.0585	0.0172	-0.1684**	-0.5923***	5.1234***	5.5544**	3.8858*	13.9560***
	(0.0415)	(0.0854)	(0.0791)	(0.1802)	(1.8842)	(2.2505)	(2.3124)	(4.1365)
Control Y-mean	6	6	7	7	3339	3339	3329	3329
% Effect (Coef/Control Mean)	-0.909	0.267	-2.575	-9.060	0.153	0.166	0.117	0.419
Adj. R-Squared	0.053	0.018	0.110	0.034	0.269	0.085	0.457	0.135
Obs	440037	440037	71604	71604	440037	440037	71604	71604
C,Year FE	Y	Y	Y	Y	Y	Y	Y	Y
1970 controls-pscore weighting	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Table reports the estimated coefficients and standard errors using the Sun and Abraham (2021) method, controlling for county and year of the second trimester. Panel A presents results for Black mothers, and Panel B for White mothers. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells. While Columns 1 and 5 show estimates weighted by the propensity score multiplied by the cell size, Columns 2 and 6 report estimates from separate regressions using propensity score weighting without multiplying by the cell size. Columns 3 and 7 and Columns 4 and 8 repeat the same exercise for cells that have at least 2 black and white births in a given month and year of the second trimester. Standard errors are clustered at the county level. $*\ p<0.10$, $**\ p<0.05$, $***\ p<0.01$.

Table A.11: Robustness to Alternative Timing of Treatment Assignment using Individual-level Data: The Effect of WIC on Low Birth Weight and Birth Weight

	Black Mothers					White Mothers				
	Pref: 30.5 days, 40 wks. to miss,	30 days, 40 wks. to miss.	30.5 days, 39 wks. to miss.	30.5 days, . to miss.	Pref: 30.5 days, 40 wks. to miss.	Pref: 30.5 days, 40 wks. to miss.	30 days, 40 wks. to miss.	30.5 days, 39 wks. to miss.	30.5 days, . to miss.	Pref: 30.5 days, 40 wks. to miss.
A: Fraction below 2,500 Grams										
WIC-month year-trim2	-0.3078***					-0.0585				
	(0.1095)					(0.0415)				
WIC-month year-trim2		-0.3070***					-0.0601			
		(0.1128)					(0.0420)			
WIC-month year-trim2			-0.3078***					-0.0585		
			(0.1095)					(0.0415)		
WIC-month year-trim2				-0.3513**					-0.0565	
				(0.1738)					(0.0522)	
WIC-month year-trim3					-0.3230***					-0.0552
					(0.1161)					(0.0406)
Y-mean	13	13	13	13	13	6	6	6	6	6
Adj. R-Squared	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.002	0.002
Obs	5735201	5743132	5735201	3517391	5626681	33089428	33142458	33089428	22413097	32419034
B: Birth Weight (in Grams)										
WIC-month year-trim2	11.7089***					5.1234***				
	(4.0134)					(1.8836)				
WIC-month year-trim2		11.5870***					5.1313***			
		(3.9914)					(1.9014)			
WIC-month year-trim2			11.7089***					5.1234***		
			(4.0134)					(1.8836)		
WIC-month year-trim2				13.8598**					5.3635**	
				(5.5453)					(2.2986)	
WIC-month year-trim3					11.3854***					5.4370***
					(3.9841)					(1.8786)
Y-mean	3080.089	3080.210	3080.089	3096.427	3079.654	3339.085	3339.259	3339.085	3352.783	3338.197
Adj. R-Squared	0.005	0.005	0.005	0.006	0.005	0.009	0.009	0.009	0.010	0.009
Obs	5735201	5743132	5735201	3517391	5626681	33089428	33142458	33089428	22413097	32419034
C,Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
1970 controls -pscore weighting	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Table reports the estimated coefficients and standard errors using the Sun and Abraham (2021) method, controlling for county and year of the second trimester with an estimate of the propensity of WIC adoption based on pre-program characteristics. Panel A presents results using low birth weight as the outcome, and Panel B presents results using birth weight. The analysis is run using individual-level data instead of cell-level data, alternative use of gestational age; and alternative assignment of WIC timing (month and year of the third trimester). We test robustness using alternative assignment of gestational age because for a subset of the sample, gestational age is missing. In our preferred estimation, we assign 30.5 days to a month and 40 weeks to the missing gestational age to construct the gestational age for the full sample (Columns 1 and 6) following Currie and Schwandt (2013). We use 30 as the average number of days in a month (Column 2 and 7), assign 39 weeks for missing gestational age (Columns 3 and 8), or leave gestational age as missing (Columns 4 or 9) in our sample. In Columns 5 and 10, we assign the WIC exposure in the month-year of the start of the third trimester. We use data from Vital Statistics Natality Files that include births with the year of the second or third trimester between 1968 and 1980. We describe the construction of the trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The regressions are weighted by one for the treatment group, and $\hat{p}/(1-\hat{p})$ for the control group. We compile data on WIC timing from multiple sources between 1974 and 1981. Standard errors are clustered at the county level. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

Table A.12: Robustness to Assigning WIC Exposure by "Year" rather than the Second Trimester: The Effect of WIC on Low Birth Weight and Birth Weight

	Low Birth Weight			Birth Weight		
	Year Trimester 2	Year Trimester 3	Year Birth	Year Trimester 2	Year Trimester 3	Year Birth
A: Black Mothers						
WIC-year-trim2	-0.2385*			10.8016**		
	(0.1249)			(4.2281)		
WIC-year-trim3		-0.2417*			10.2256**	
		(0.1286)			(4.2133)	
WIC-year-birth			-0.2225*			9.9266**
			(0.1252)			(4.3970)
Control Y-mean	13	13	14	3080	3079	3077
Adj. R-Squared	0.147	0.151	0.158	0.355	0.352	0.350
Obs	22870	22830	22620	22870	22830	22620
B: White Mothers						
WIC-year-trim2	-0.0483			5.4705***		
	(0.0404)			(1.8288)		
WIC-year-trim3		-0.0430			5.7416***	
		(0.0418)			(1.7960)	
WIC-year-birth			-0.0419			5.4369***
			(0.0409)			(1.8185)
Control Y-mean	6	6	7	3338	3337	3335
Adj. R-Squared	0.367	0.355	0.368	0.773	0.770	0.772
Obs	37477	37513	37438	37477	37513	37438
C,Year FE	Y	Y	Y	Y	Y	Y
1970 controls-pscore weighting	Y	Y	Y	Y	Y	Y

Notes: Table reports the estimated coefficients and standard errors using the Sun and Abraham (2021) method, controlling for county and year of the second trimester. Panel A presents results for Black mothers, and Panel B for White mothers. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race, county, and the respective year. Columns 1 and 4 show estimates that assign WIC exposure based on the year of the start of the second trimester, while Columns 2 and 5 use the year of the third trimester. Columns 3 and 6 assign exposure based on the year of birth. Standard errors are clustered at the county level. $*\ p<0.10$, $**\ p<0.05$, $***\ p<0.01$.

Table A.13: Additional Robustness to Trimming p-score and Excluding The CSFP Control: The Effect of WIC on Low Birth Weight and Birth Weight

	Low Birth Weight			Birth Weight		
	Preferred: p-score with CSFP Ever Control, No Trimming	p-score with CSFP Ever Control, With Trimming	p-score without CSFP Ever Control, No Trimming	Preferred: p-score with CSFP Ever Control, No Trimming	p-score with CSFP Ever Control, With Trimming	p-score without CSFP Ever Control, No Trimming
A: Black Mothers						
WIC-month year-trim2	-0.3078***	-0.2964**	-0.2500**	11.7089***	11.0420**	10.1285**
	(0.1095)	(0.1247)	(0.1051)	(4.0157)	(4.4079)	(4.1450)
Control Y-mean	13	13	13	3080	3080	3080
Adj. R-Squared	0.028	0.026	0.034	0.079	0.075	0.089
Obs	189417	172366	189417	189417	172366	189417
B: White Mothers						
WIC-month year-trim2	-0.0585	-0.0502	-0.0719	5.1234***	4.6284**	5.5395***
	(0.0415)	(0.0456)	(0.0506)	(1.8842)	(1.9148)	(1.9277)
Control Y-mean	6	6	6	3339	3341	3339
Adj. R-Squared	0.053	0.049	0.057	0.269	0.257	0.285
Obs	440037	410908	440037	440037	410908	440037
C,Year FE	Y	Y	Y	Y	Y	Y
1970 controls-pscore weighting	Y	Y	Y	Y	Y	Y
p-score, no trimming	Y			Y		
p-score, with trimming		Y			Y	
Control for CSFP ever	Y	Y		Y	Y	

Notes: Table reports the estimated coefficients and standard errors using the Sun and Abraham (2021) method, controlling for county and year of the second trimester with an estimate of the propensity of WIC adoption based on pre-program characteristics. Panel A presents results for Black mothers, and Panel B for White mothers. In Columns 1 and 4, we show our preferred specification, not trimming propensity score and controlling for the presence of the Commodity Supplemental Food Program (CSFP), which provided nutritional assistance to low-income families with pregnant women or children under 6. In Columns 2 and 5, we report the results after trimming the propensity score by dropping counties with more than 0.99 probability of having WIC to address potential concerns about the counties with a high propensity of having WIC affecting the results. In Columns 3 and 6, we present our results without the presence of CSFP. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells. Standard errors are clustered at the county level. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

Table A.14: Robustness to Dropping Counties Serving Special Population (Migrants and Native Americans): The Effect of WIC on Low Birth Weight and Birth Weight

	Low Birth Weight		Birth Weight	
	Preferred- All Programs Regular and Spec. Pop.	Drop counties Special Pop. Only	Preferred- All Programs Regular and Spec. Pop.	Drop counties Special Pop. Only
A: Black Mothers				
WIC-month year-trim2	-0.3078***	-0.3110***	11.7089***	11.7325***
	(0.1095)	(0.1098)	(4.0157)	(4.0262)
Control Y-mean	13	13	3080	3080
Adj. R-Squared	0.028	0.028	0.079	0.079
Obs	189417	187283	189417	187283
B: White Mothers				
WIC-month year-trim2	-0.0585	-0.0599	5.1234***	5.1710***
	(0.0415)	(0.0418)	(1.8842)	(1.9013)
Control Y-mean	6	6	3339	3338
Adj. R-Squared	0.053	0.053	0.269	0.269
Obs	440037	429441	440037	429441
C,Year FE	Y	Y	Y	Y
1970 controls-pscore weighting	Y	Y	Y	Y

Notes: Table reports the estimated coefficients and standard errors using the Sun and Abraham (2021) method, controlling for county and year of the second trimester. Panel A presents results for Black mothers, and Panel B for White mothers. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980 based on pre-program county characteristics. The cells are weighted by the number of births in the cell for the treatment group, and $\hat{p}/(1-\hat{p})$ times the number of births in the cell for the control group. We use data from Vital Statistics Natality Files that include births with the year of the second trimester between 1968 and 1980, corresponding to infants born between February 1968 and September 1981. We describe the construction of the second trimester timing in Section 3. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level by race-county-year-of-second-trimester cells. Standard errors are clustered at the county level. $*\ p<0.10$, $**\ p<0.05$, $***\ p<0.01$.

Table A.15: Replication of Hoynes et al. (2011) (HPS)

	All HPS counties HPS rollout		HPS counties no VA HI HPS rollout		HPS counties no VA HI OUR rollout		OUR counties OUR rollout	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Fraction below 2,500 Grams								
WIC exposure HPS	-0.0053	0.0099	-0.0108	0.0068				
	(0.0384)	(0.0407)	(0.0383)	(0.0407)				
WIC exposure US					-0.0062	0.0043	-0.0249	-0.0075
					(0.0359)	(0.0386)	(0.0274)	(0.0295)
Y-mean	6.795	6.795	6.777	6.777	6.795	6.795	6.722	6.722
R-Squared	0.753	0.763	0.754	0.765	0.750	0.760	0.723	0.731
Obs	18598	18598	18274	18274	19013	19013	29018	29018
N Counties	1996	1996	1917	1917	1917	1917	2940	2940
B: Birth Weight (in Grams)								
WIC exposure HPS	2.3444**	1.8838*	2.5933**	2.0080*				
	(1.1712)	(1.1324)	(1.1714)	(1.1319)				
WIC exposure US					2.1604*	1.7810*	3.0458***	2.2554***
					(1.1741)	(1.0559)	(0.8851)	(0.8605)
Y-mean	3342	3342	3343	3343	3342	3342	3345	3345
R-Squared	0.901	0.909	0.902	0.910	0.901	0.908	0.889	0.896
Obs	18598	18598	18274	18274	19013	19013	29018	29018
N Counties	1996	1996	1917	1917	1917	1917	2940	2940
C,Y FE	Y	Y	Y	Y	Y	Y	Y	Y
REIS controls+FS exposure	Y	Y	Y	Y	Y	Y	Y	Y
1970 controls*trend	Y	Y	Y	Y	Y	Y	Y	Y
State*Y FE	N	Y	N	Y	N	Y	N	Y

Notes: Hoynes et al. (2011) run their analysis using cell-level data collapsed to the "county and year of the third-trimester" level for 1971-1975 and 1978-1982 and TWFE specification. WIC exposure is assigned in the year of the third trimester. The sample excludes cells that are less than 25 births. In addition, we drop Virginia from their analysis. REIS controls+FS exposure include controls for per capita government transfers, including Social Security retirement and disability payments, Medicare and Champus spending, cash public assistance (AFDC and SSI), income per capita, and Food Stamp exposure. 1970 county controls include the percentage of population that is non-White, poor, urban, young, and old age, and the natural log of population. Columns 1 and 2 use their original rollout, and all their counties in their sample. Columns 3 and 4 use their rollout but exclude counties in Virginia and Hawaii. Columns 5 and 6 use our rollout variable for their county sample. Columns 7 and 8 use our full rollout variable that added WIC information for a sizable number of counties. Estimates are weighted using births in the cell. Standard errors are clustered at the county level. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

Table A.16: Joint Significance of Point Estimates, with and without Multiple Hypothesis Correction

		Family Wise Error Rate	False Discovery Rate
Outcome	p**-values from regressions (1)**	p**-values by Holm (1979) (2)**	q**-values by Benjamini et al. (2006) (3)**
Panel A: Black Mothers			
Low Birth Weight	0.005	0.035	0.014
Very Low Birth Weight	0.005	0.035	0.014
Small for Gestational Age	0.037	0.150	0.031
Preterm	0.124	0.199	0.064
Birth Weight	0.004	0.029	0.014
High Birth Weight	0.073	0.199	0.044
Large for Gestational Age	0.008	0.041	0.014
Kessner Index for Adequate Prenatal Care	0.066	0.199	0.044
Panel B: White Mothers			
Low Birth Weight	0.159	0.952	0.496
Very Low Birth Weight	0.509	1.000	0.737
Small for Gestational Age	0.303	1.000	0.571
Preterm	0.765	1.000	0.825
Birth Weight	0.007	0.053	0.056
High Birth Weight	0.226	1.000	0.513
Large for Gestational Age	0.083	0.580	0.409
Kessner Index for Adequate Prenatal Care	0.613	1.000	0.779

Notes: Column 1 of this table reports the p-values from Table 1. In Column 2, we report p-values for a joint significance test of all pre-program coefficients after correcting for multiple hypothesis testing (MHT) using a step-down procedure introduced by Holm (1979) that controls for the family-wise error rate. In Column 3, we report p-values that adjust for MHT using the Benjamini et al. (2006) procedure that accounts for the false discovery rate based on sharpened q-values.