# Effects of WIC on Birth Outcomes: Evidence from Newly Digitized Data from the National Archives

Marianne Bitler, UC Davis & NBER
Danea Horn, Stanford University
Esra Kose, UC Merced
Maria Rosales-Rueda, University of Delaware & NBER
Arian Seifoddini, Federal Trade Commission*

November 2, 2024

**Abstract**

This paper examines the effect of nutritional investments for low-income families through the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) on infant health. We substantially extend the previous work by Hoynes et al. (2011) by incorporating new data on WIC roll-out from the National Archives, using new methods, and examining additional infant outcomes. We find robust evidence that access to WIC in a women's county of residence by the start of their second trimester led to higher birth weight and a reduction in the probability of low birth weight, with larger effects for Black mothers. Our results suggest treatment-on-the-treated effects for birth weight of around 36 grams and for low birth weight of -1.02 percentage points for Black mothers, which are both economically significant. The results for White mothers are much smaller and statistically insignificant. These findings show WIC reduced overall disparities between Blacks and Whites in low birth weight by 4.6 percent.

*Thanks to Janet Currie, Hilary Hoynes, Margaret McConnell, Katherine Meckel, Jenna Nobles, Laura Wherry, and conference participants at the NBER Children's meeting, ASHE, APPAM, HERO, and the VEPPS seminar for helpful comments, and to Hilary Hoynes, Marianne Page, and Ann Stevens for generously sharing data from their paper. 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 the U.S. government.

# 1 Introduction

Health disparities among racial and ethnic groups are an important concern to researchers and policymakers. These disparities begin at birth and persist throughout the lifetime. This is paramount 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 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).

WIC is a means-tested program, currently offering benefits to children aged 0-4 and pregnant/postpartum women whose income is below 185% of the Federal Poverty Guideline. 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 women and/or children under 4 (and, later, 5) in the household, that this person is at nutritional risk, and that they participated in or are 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 satisfied the means testing.[1]

WIC has the potential to influence health at birth and beyond, and to ameliorate gaps in nutrition-related health outcomes. First and foremost, it provides nutritious food to low-income, nutritionally at-risk postpartum, pregnant, and breastfeeding women, infants, and children under age five. Bendick Jr. et al. (1976) report that in 1976, the WIC food package for pregnant women consisted of milk or cheese, eggs, high-iron breakfast cereal, and high-vitamin fruit juice, with average monthly food costs of about $20 per month, or about 39 percent of average food at home spending in 1974 (Manchester, 1976).[2] Second, WIC also provides referrals to other health and social safety net programs and nutrition education. At the time of WIC's creation, foods were obtained in three ways. In some locations, participants obtained 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 those items up at local warehouses or, rarely, at the clinics themselves. Today, the program has transitioned to electronic benefit transfer (EBT). The research literature on WIC has established that the current incarnation of this program

[1] Later on, more explicit rules about family income being below a particular threshold were put in place.

[2] Important parts of the package for infants were infant cereal, and, in particular, iron-fortified infant formula.

improves the quality of nutritional intake and birth outcomes (e.g., Bitler and Currie, 2005; Figlio et al., 2009; Rossin-Slater, 2013; Currie and Rajani, 2015; Currie and Rossin-Slater, 2020).

We add to this rich literature by exploring the roll-out of the WIC program during the 1970s and 1980s. The WIC program was created at the tail end of the War on Poverty, with the first clinic opening in January 1974 in Pineville, KY. Over the next few years, the USDA enacted a series of program changes but was also constrained by lawsuits to expand the program significantly. We take advantage of this quasi-experimental variation in WIC implementation to study its effects on pregnancy and birth outcomes. We pay particular attention to how the program influences birth outcomes by race, focusing primarily on Black-White differences.[3]

Our paper revisits prior work by Hoynes et al. (2011)[4] who examined the effect of WIC adoption on birth outcomes using two-way fixed effects (TWFE) models. We have augmented their roll-out data with extensive new data from the National Archives on State WIC plans through 1980, allowing us to detect with more granularity the time and geographic coverage of WIC introduction and to uncover nearly 450 additional counties that had WIC present. We also use new methods that are robust to treatment effect heterogeneity across places and over time, as well as to contamination from other periods' effects. In addition, we focus on examining differential effects by race and explore mechanisms that could explain that heterogeneity.[5]

Our preferred estimates use inverse propensity score weights to adjust the control group observations so that the pre-WIC cross-county distribution of economic conditions, demographics, characteristics of the health care system, other War on Poverty Programs, and birth outcomes are balanced across the treatment and control groups. Our preferred methods use difference-in-differences combining propensity scores with Sun and Abraham (2021)'s approach. The results indicate that the adoption of the WIC program is associated with a 0.37 percentage point reduction in the incidence of low birth weight among Black infants, along with a 12.8 gram increase in average birth weight, both estimates are statistically

[3]We exclude Native Americans from this study because Native Americans who resided on Indian Land Areas often had access to WIC run by Indian Tribal Organizations, which may have been quite different 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]We use "HPS" to reference this important work.

[5]Malinovskaya (2023) looks at long-run effects of WIC exposure in utero on adults, using a combination of the HPS WIC roll-out and more detailed information for the state of 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.

significant and economically meaningful. This compares to a statistically significant increase in average birth weight of 4.9 grams for White infants and a statistically insignificant decline of 0.05 percentage points for White infants. These results suggest that WIC led to a decline of around 8 grams on average in the overall Black-White birth weight gap, or 3 percent, and a decline of about 0.32 percentage points, or 4.6 percent, in the low birth weight gap. We further investigate the impacts of WIC at different points in the birth weight distribution and find that the effects are larger at the lower end, mainly for Black infants.

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). We consider multiple reasons that might plausibly explain why WIC improves infant health outcomes for Black mothers more than for White mothers. First, eligibility, participation, and take-up of WIC were substantially higher for Blacks than Whites. Second, Black infants in this period had far worse baseline birth outcomes than their White counterparts.[6] Third, individual risk factors (e.g., teen pregnancy and single parenthood) affect Black mothers and White mothers in different ways. Fourth, differences in place-based factors faced by Black mothers versus 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 results may be explained by (1) baseline differences in outcomes by race, leading to higher returns on investments for Black mothers relative to White mothers, and (2) racial differences in WIC program take-up and eligibility.

## 2 The WIC program and WIC roll-out

### Background

The WIC program, at its inception, provided access to either vouchers to obtain healthful foods rich in nutrients of concern or to home-delivered food or access to food at warehouses or clinics to pregnant, postpartum, and breastfeeding women and young children using or eligible for free or subsidized health care, who were also nutritionally at risk. Participants also had to live in the "area served" by the relevant local WIC program.

[6] Appendix Figure A.1 shows that Black infants had a worse baseline birth weight in 1970 than Whites.

Relevant to our study, WIC could provide a significant share of essential foods and nutrients for pregnant women. For example, in 1976, in California, WIC benefits for pregnant women covered 60–100% of the required dietary allowance for protein, calcium, iron, vitamin C, and vitamin A, contributing around 25% of calories. These nutrients have been identified as critical for the healthy development of the fetus during pregnancy (Shapira, 2008; Marshall et al., 2022).

## Roll-out of WIC between 1974 and 1980

The WIC program was established by law in 1972, but the first clinic became operational in early 1974. A series of lawsuits from outside groups and efforts by the Nixon and Ford administrations to shrink, change, or end the WIC program led to considerable variation over time and across places when 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 roll-out. Numerous congressional hearings include presentations from non-profits who wanted to establish WIC clinics but were not selected. This is important for establishing our key assumption that the introduction and location of WIC clinics is idiosyncratic, conditional on the pre-WIC characteristics of counties. Below, we explore the extent to which the timing and location of WIC clinics are explained by counties' economic or demographic conditions or the presence of the other War on Poverty programs. Also, we account for pre-WIC characteristics in our analysis.

We build upon WIC roll-out 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 Special Supplemental Food Programs for Women, Infants and Children.

We fill in many of the Hoynes et al. (2011) data gaps using National Archives sources and some other documents.[7] We digitized information from WIC Program State plans that

[7] Hoynes et al. (2011) have limited data for 1976 and 1977, but we also have less granularity that year. 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 earlier than FY 1978. They were implemented because the USDA lost a lawsuit and was forced to both open the program to communities with needy applicants and document this in the State Plans.

states had to provide to the USDA from fiscal year 1978[8] to 1981.[9] In 1976, a court order (based on a lawsuit that the USDA lost) required the agency to ensure that the program was extended to those women and infants most in need. As a result, the USDA required State agencies to develop Affirmative Action Plans as part of the WIC program's State Plans, which provided us access to WIC implementation data.

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 concerning special populations served (migrants, Native Americans). Also, the specific information included and how it was presented varied by state. We digitized these State Plans and some other documents to measure roll-out. With our newly digitized data combined with the digitization of the HPS sources, we create a measure of WIC roll-out 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 and administrative caseload 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. In 1974, WIC started in 356 counties, and its adoption expanded significantly throughout the rest of the 1970s, with 2,376 counties having adopted by 1980.[10] Panel (b) of Figure 1 presents the geographic variation in WIC adoption timing, which shows that there is a lot of variation across and within states across counties. We take advantage of this variation to examine the effect of WIC on infant health outcomes.[11]

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

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

[10] We exclude Virginia, Alaska, and Hawaii from our sample due to missing county characteristics (HI) and changing geography (VA, AK). We also do not have comprehensive roll-out documents in 1976 and 1977, and the number of counties in those years comes from retrospective information in some of the State WIC plans. We note that other documents show that roll-out in those years was relatively rare (Council et al., 1982).

[11] Figure A.2 compares our compilation of WIC roll-out with that from Hoynes et al. (2011), showing we found a substantive share of new roll-out locations (around 450 additional counties) relative to this preexisting literature due to our novel sources for roll-out. In addition, Table A.1 reports cross-tabulations of the timing of adoption for 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.

## 3 Data

Data on birth outcomes are from the Vital Statistics birth records from 1968 to 1981.[12] These data are created 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 birth weight and whether the infant is considered low birth weight (birth weight $< 2,500$ grams, an important threshold at the time). We also examine several additional outcomes from the birth certificate: whether the infant was classified as having a very low birth weight (birth weight $< 1,500$ grams), having a very high birth weight (birth weight $> 4,500$ grams), and whether the infant was small for gestational age or large for gestational age;[13] and whether the infant was born prematurely (before 37 weeks of gestation). The prenatal care measure we use is constructed using the Kessner index for adequate prenatal care.[14] Lastly, we explore the birth weight distribution in greater depth.

The birth certificate data also includes information on the mother's age and race, child gender, and, in some states and years, marital status, place of residence of the mother, and mother's educational attainment. More details about our sample selection are available in the 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. Our results focus primarily on Black and White infants, 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 some pre-WIC county characteristics across the treatment and control groups as discussed in Section 4 in more detail. 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

[12] Vital Statistics micro-data are available starting in 1968, and county identifiers are publicly available through 1988. We stopped with 1981 births/1980 second trimesters because this was the end of our digitized data on WIC adoption timing, and there were also program cutbacks in FY 1981 due to Reagan Administration policy choices. At the end of the period, we also have novel data on participation and take-up by county in 1980, but there is nothing more comprehensive for our intervention period.

[13] "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.

[14] Kessner index for adequate 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.

(DD) models we use, which allow for heterogeneity in effects across counties or over time by race group. However, our results are robust to using traditional TWFE models, individual-level data, and diverse sets of controls and specifications. Note that 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).

We combine these Vital Statistics data with the WIC roll-out and various other county-level characteristics, including the county population in 1970; exposure to other important War on Poverty programs shown to affect birth outcomes (Food Stamps, Community Health Centers); and economic, safety net, and demographic characteristics (such as spending on the Medicare program and unemployment insurance compensation, a proxy for the county level labor market).

As mentioned above, we assign exposure to WIC based on the calendar month and year of the start of the second trimester of the birth, using the reported gestational age where available and using 40 weeks when gestational age is missing. We do so because 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). Around 30% of births have missing gestational age. For these individuals, we assume 40 weeks of gestation (the median).[15] After calculating the month and year of conception, we calculate the month and year of the start of the second trimester. It is worth noting that our results are 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.[16] For time-varying covariates measured at the year level, we match them according to the year of the second trimester.

We present summary statistics for outcome and control variables in Appendix Table A.2. On average, 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 higher access to Food Stamp 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

[15]Note that 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.

[16]We discuss these results in Section 5.4 and present these results in Appendix Table A.10.

non-WIC areas. 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 baseline disadvantage, in line with the human capital theory.

## 4 Empirical Strategies

We exploit temporal and geographic variation in the roll-out 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 as the control group (counties adopting WIC after 1980 in our context). As explained below in detail, to address the issue that never-treated counties may be different from those that adopted WIC during our study period, we use a propensity score reweighting approach related to Sant'Anna and Zhao (2020)'s doubly robust approach.

The standard difference-in-difference specification is:

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

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.[17] $C_{ct}$ are county time-varying controls that could confound the effects of WIC and include the 1970 population as well as hospitals per capita, the presence of Food Stamps, and Community Health Centers in the county. We account for these control covariates in two ways: our preferred specification includes them as pre-program covariates in the computation of the propensity score weights, while we also estimate specifications that directly control for them as robustness checks. $\gamma_c$, $\tau_t$ are county and year of the start of the second trimester fixed effects, which absorb time-invariant county characteristics that could be correlated with WIC roll-out and time-varying national changes, respectively. Each cell observation is weighted by the number of births in the cell, and our estimates generally are stratified by race, with a focus on Black and White mothers.[18] Standard errors are clustered at the county level. The coefficient of interest is $\beta$, which measures the static DD estimate

[17] Our results are robust to defining exposure according to the start of the third trimester of pregnancy.
[18] We also present the results for non-Native "Other" mothers.

of the effect of WIC. This coefficient is identified from changes in birth outcomes between births exposed to WIC and non-exposed before and after the roll-out.

In our setting, we are concerned with treatment effect heterogeneity across counties adopting in the same month/year of the second trimester and across counties adopting in different months/years of the second trimester, which motivated us to implement the approach of Sun and Abraham (2021). Our control group in the Sun and Abraham (2021) is the set of counties that have not adopted WIC by 1980. 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 roll-out across counties is uncorrelated with other county, time-varying determinants of our outcomes of interest. To provide evidence on this assumption, we test whether 1970 county characteristics (pre-program) predict the timing of a county's WIC program implementation in Table A.3, 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 the year of WIC's introduction, expressed as an index equal to one beginning in 1974. The results, presented in Table A.3, include models with and without state-fixed effects, with results that are both unweighted and weighted by county number of births in 1970. We find that urbanicity/rurality, population, poverty level, the number of hospitals per capita, the number of active medical doctors per capita, and the presence of a Community Health Center (CHC) before 1975 are all associated with earlier WIC introduction. However, the quantitative importance of these predictors is small, and most of the variation remains unexplained as illustrated by the low R-squared, which ranges between 0.1 and 0.4.

It is also important to note that despite the requirement to create an affirmative action plan reported in the State Plan to consider where to expand WIC after the USDA lost a lawsuit in 1976, there is no evidence that the share of non-White persons in the county or the infant mortality rate predicts timing in our sample period (the coefficients are quite small and statistically insignificant). This supports our identification strategy as there is no glaring evidence of endogeneity in the timing of WIC roll-out. However, unobservables may still be a concern. To ameliorate this issue, we take two approaches, which are explained in detail below.

## 4.2 Propensity-score weights

First, we note that any DD approach relies on our adopting counties being comparable to our non-adopting counties, which may be challenging in our setting as WIC rollout is not random. Table A.4 Column 1 shows strong pre-treatment characteristics imbalances between WIC and no-WIC (by 1980) counties. 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 estimates use an alternative approach related to but simpler than Sant'Anna and Zhao (2020)'s doubly robust approach and that of Cunningham and Goodman-Bacon (2024) by using a propensity score that predicts whether a county had adopted WIC by 1980 as a function of pre-program county characteristics (mostly measured in 1970) and whether the county had other War on Poverty programs pre-WIC implementation. This approach helps to construct a valid comparison group of never-treated counties.

We adjust the weights of possible control county cells by the counties' probability of being treated based on pre-1974 county characteristics. 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 Community Health Centers and the presence of Food Stamps, county demographics, measures of the breadth of the health care system, as well as 1970 poverty and population. More precisely, the model includes measures used to designate counties as medically under-served in the 1980s (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 population controls such as the share of the 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, State Unemployment Insurance payments, and the 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, and state-level measures of White and Black AFDC participation by women and children (the latter used by (Goodman-Bacon, 2021) as a proxy for Medicaid eligibility for those on AFDC). Figure A.6 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 strong overlap.

After estimating the predicted probability of being treated $\hat{p}$, we compute cell weights as follows. For treated counties, the cell weight is left as 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.[19] This is our preferred specification because it ensures our counties are similar along all these dimensions before WIC was implemented without controlling for a host of other county time-varying controls, which can contaminate the ATE estimates. To corroborate this, Table A.4 Column 4 presents balance tests of pre-program characteristics after re-weighting the comparison counties. Most of the imbalances disappeared, and the differences are not statistically significant after adjusting for multiple hypotheses testing. Similarly, Column 7 shows that within states, there are no imbalances (except for the percent first born in 1970) between WIC-treated and never-treated counties.

We test the robustness of our results with and without controls for county Community Health Center exposure, population, and the number of hospitals per capita each year, as well as the Food Stamp Program roll-out. We present these results in Setion 5.4 paragraph one. Importantly, our conclusions do not rely on the inclusion of these covariates. In summary, these estimates are very similar to those from our preferred specification using Sun and Abraham (2021) with propensity score reweighing and those from other specifications that only include county and year-fixed effects, as we show in the 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 state-by-year fixed effects.

## 4.3 Parallel trends with p-score weighting

The second key assumption 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.

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. Also, we formally test the presence of pre-trends by computing the F-test of joint significance of all pre-treatment coefficients separately for Black and White mothers. The p-values reported inside the figures indicate that we cannot reject the joint hypothesis that all the pre-period coefficients are zero for most of the outcomes for Black mothers. Thus, we do not find strong evidence of pre-trends for Black mothers. However, for White mothers, p-values are less than 0.05 for half of the outcomes, so we are more

[19] Note that 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.

cautious about the estimated effects for White mothers.[20] In addition, we perform the pre-trends analyses developed by Roth (2022). 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.[21] 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 and Abraham, 2021; Callaway and Sant'Anna, 2020; de Chaisemartin and D'Haultfœuille, 2020). These studies have shown that causal identification under TWFE relies on the assumption of homogeneous treatment effects across location and year of adoption of the policy, which fails if treatment effects vary on average across cohorts (groups of counties that adopted the policy in the same year) or across time. The bias here can arise because earlier-treated units are used as comparison groups for later-treated units, which is only valid in the case of constant treatment effects, and the situation is more complicated with time-varying controls and groups of different sizes across time. In the presence of such heterogeneous average treatment effects by time of adoption, this comparison may lead to bias since the parallel trends assumption may no longer hold for some of the comparisons of later-adopting locations to already-adopted locations in the presence of non-constant effects.

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). These authors show that the TWFE estimator is a weighted sum of several difference-in-differences estimators across cohorts and periods.[22] These weights sum to one, and some may be negative. Their proposed test first estimates the standard TWFE model, extracts the implied weights, and then examines whether these estimated weights are correlated with factors that

[20] Additionally, we report the p-values from a joint significance test for all the pre-period coefficients before (Column 1) and after adjusting for multiple hypothesis testing (Column 2) in Table A.5. This additional analysis shows that there are statistically significant pre-trends for only one outcome for Black mothers (adequate prenatal care) and at least two outcomes for White mothers (low birth weight, adequate prenatal care) at a 5% significant level.

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

[22] These difference-in-differences estimators can be of the following types: For early treated groups using never treated as a comparison, for later treated groups using never treated as a comparison, for early treated groups using later treated as a comparison, and for later treated groups using early treated as a comparison. The last type is particularly problematic as early treated groups may not be a valid comparison group for later treated groups in the presence of treatment effect heterogeneity (Goodman-Bacon, 2021).

are associated with the size or the intensity of the treatment effects. In our context, these test results are presented in Appendix Table A.6.[23] 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 15% of the TWFE weights are negative. This evidence supports that TWFE may not be appropriate in our setting because heterogeneous treatment effects can lead to bias.

For this reason, we use the procedure developed by Sun and Abraham (2021) (SA) that accounts for dynamic treatment effects across 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 who 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). Then, to calculate the static DD estimator, the estimated CATTs are averaged across the post-period using the sample share of each cohort as weights. This method gives a consistent estimate of the treatment effect under the parallel trends and no anticipation assumptions. Sun and Abraham (2021) show that their estimator is robust to heterogeneous treatment effects. For comparison, we present estimations using TWFE in the robustness section. We note that in our setting, the "never treated group" is the set of counties that had not adopted WIC by the fiscal year 1980.

# 5 Results

## 5.1 Infant health outcomes

Table 1 presents our static effects of WIC on several birth outcomes using Sun and Abraham and the propensity score reweighting approach. These outcomes include low birth weight (<2500 grams), very low birth weight (<1500 grams), small for gestational age (below the tenth percentile of birth weight for gestational age), preterm birth (<37 weeks gestation), birth weight (in grams), high birth weight (>4500), large for gestational age (above the ninetieth percentile of birth weight for gestational age), and Kessner index for adequate prenatal care (constructed based on gestational age and the number of visits) in Columns 1 to 8, respectively.

These results show that the intent-to-treat (ITT) effects of WIC's introduction in a county

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

are significant and meaningful for Black births relative to White births. We find that the improvements in health at birth measured by the probability of low birth weight, very low birth weight, and small for gestational age are larger among Black births than among White births (Columns 1-3). Black births exposed to WIC in their county experienced a statistically significant decline in low birth weight of 0.37 percentage points (Column 1), compared to a decline of 0.05 percentage points for White births. Relative to the pre-period mean, the results imply ITT effects of -2.8% and -0.75%, 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 5.2% and 5%, respectively, for Black babies compared to 0.9% and 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).

In Column 4, the results show that WIC exposure increases the likelihood of preterm birth among Black babies (statistically significant at the 10% level). This result is puzzling. We note, however, that the determinants of preterm are largely unknown (e.g., Butler and Behrman, 2007).

Turning to birth weight in Column 5, WIC led to a statistically significant increase in this marker by 13 grams for Black babies compared to 5 grams for White babies. Also, we find that WIC exposure increases the probability of high birth weight (Column 6) and large for gestational age (Column 7), particularly for Black births. However, we note that these outcomes are quite rare. These findings suggest that as a nutritional intervention, WIC, not only affected birth weight on the left tail (low birth weight) but also moved some births from normal birth weight to 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 care, particularly for Black births.

[Add a paragraph or a footnote about Infant Mortality results in this section?]

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) suggestive evidence of dynamic effects up to 3 years in the post period and (2) no strong evidence of pre-trends in most outcomes. Importantly, the figures show a gradual decline in the likelihood 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 small and noisy, and pre-trends are somewhat more present.

We further assess the parallel trends assumption and potential violations using a couple of tools. First, we formally test the presence of pre-trends using a joint F-test for the significance of all the pre-treatment coefficients. For Blacks, p-values reported in the figures indicate that two out of eight variables show the presence of pre-trend at the 5% significance level, namely high birth weight and Kessner index for adequate prenatal care. In addition, small for gestational age and average birth weight pre-trends are significant at the 10% level. Additionally, Appendix Table A.5 presents these p-values adjusted for multiple hypothesis testing (MHT), which shows that only one outcome (adequate prenatal care) remains significant. Overall, these results indicate no strong evidence of pre-trends in most of the outcomes for Black mothers. For Whites, p-values reported in the event study figures reveal the potential presence of pre-existing trends in some birth outcomes (and two of them remain significant after MHT adjustments). Therefore, the results for White mothers should be taken with caution.

Second, following Roth (2022), to explore how well-powered our event study pre-trends are, we estimate the slope of a linear pre-trend that we could detect with 80 percent power. Then, based on that linear pre-trend, we estimate a counterfactual evolution for each outcome, which we include in the event-study figures. This analysis suggests that for Blacks, the post-WIC estimates (including their confidence intervals) are larger than the counterfactual trend for low birth weight and small for gestational age, further validating our results for these outcomes. For Whites, post-WIC estimates are noisier and insignificant and overlap with the counterfactual trend.

### 5.1.1 Treatment-on-the-treated (TOT)

To translate the ITT estimates into average TOT effects, we need information on the fraction of pregnant women who participated in WIC by race during our study period. These are reported in Figure 2.[24] To approximate that information, we use the only information available, to the best of our knowledge, about WIC participation by race, from 1986 from Food Research and Action Center (1988). The WIC participation rate (the ratio of the number of Black participants who are women to the number of singleton births in a given

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

year) was 36% for Black pregnant women and 15% for White women.[25] However, these participation numbers are likely 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 at most -1.02 percentage points for Black babies and at most 0.32 percentage points for Whites (-7.5% and -5% of the control mean, respectively).

## 5.2 Distributional effects for birth weight

We explore further heterogeneity in the effects of WIC on birth weight by looking at various cuts across the birth weight distribution, namely birth weight below 1500, 2000, 3000, 3250, 3500, 3750, 4000 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 Black babies (in lavender color) and White babies (in grey color). For example, 13.2 percent of Black babies had a birth weight under 2500 grams (low birth weight), while only 6.2 percent of White babies did. These baseline means show that the White advantage in birth weight is present across the birth weight distribution.

Overall, we find that WIC shifts the 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 and low birth weight, WIC decreased the share of Black births with birth weights under 2000 grams, 3000 grams, 3250 grams, and 3500 grams (normal birth weight), as well as 3750 grams, and 4000 grams. It led to a small increase in the probability of being a high birth weight/reduction in being below 4500 grams. For each of these measures, the White estimates are much smaller in magnitude. For the share of birth weight under 1500 grams, 2000 grams, and 2500 grams, the equality of estimates among races can likely be rejected at the 90% confidence level, and for some, we may even reject it at the 95% confidence level.

Panel (b) converts these estimates into percent effects by dividing the point estimates with the baseline means. For Black babies, the effects (in percent terms) are larger for the lower end of the birth weight distribution. Once again, the percent effects are larger in

[25] 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 would be to divide by something larger than these participation numbers because 36% average participation for Black mothers represents 0 participation in some places and participation above 36% in others. We would ideally use the highest percent participation.

magnitude for Blacks than Whites. This finding suggests that the racial heterogeneity in WIC effects is not simply driven by the differences in baseline means for Black and White births.

## 5.3 Potential explanations underlying the race heterogeneity impacts

Why does WIC improve infant health outcomes for Black mothers more than for White mothers, even after adjusting for differences in the participation rate?[26] In particular, we show that the approximated TOT effects on low birth weight are likely larger in magnitude for Blacks.[27] To answer this question, we explored six possible explanations. For five of the six, we stratify our samples by either the individual or baseline county/state explanatory factors (or other factors) and estimate the effects of WIC within each sample.

The first explanation we consider is the role of eligibility in explaining our findings. Panel (a) of Figure 2 shows the average state-level eligibility by race for the program, calculated using the March Current Population Survey (CPS) for 1988–1990 for 0-year-olds in addition to participation by race used above. We note that the March CPS collects income and program participation data for the previous calendar year. Thus, those who are 0 years old in March were likely born in the previous calendar year, and we can use the income and program participation levels of their household in the previous calendar year to approximate eligibility by race.[28] We then summarize this by race for the 0-year-olds, constructing a weighted sum using the CPS person weights for the infants. Blacks have a much higher eligibility rate than do Whites, with 74% of Black infants likely having been eligible compared to 36% of Whites.

A second possible explanation is that take-up rates are higher for Blacks than Whites. We explore this by combining the participation numbers by race with the share of WIC-eligible individuals calculated before (ratio of numbers in Panel (b) to Panel (a) of Figure 2 for each

[26] 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 statistically significant result and is quite large (21 grams improvements following WIC's introduction) relative to the estimated effects for Blacks (12.7 grams) and Whites (4.9 grams). These results are reported in Appendix Table A.7.

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

[28] We determine eligibility as being in a family whose income is under 185% of the federal poverty guideline or were participating in Food Stamps, the National School Lunch Program, AFDC, or Medicaid. Food Stamps, AFDC, and Medicaid confer automatic, also known as adjunctive, eligibility for the WIC program by 1989. The School Meals shared the same eligibility threshold and also conferred eligibility in some places, while the other 3 programs had thresholds below the maximum WIC threshold.

race). Note that the denominators of both numbers are singleton births, and thus, this ratio is the number of participants per eligible person or takeup. This suggests a take-up rate among eligibles of about 47.4 % for Blacks, compared to one of 44% for Whites. Together, the first two explanations suggest a limited role for take-up differences by race but a possible role for eligibility in explaining part of these differences. We discuss what we find with a 1-year take-up measure below.

A third explanation we explored is the level of individual risk. We first 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 on the birth rate (Column 1), on the log of the number of births, on the share with teen mothers, or mothers with low or high education. We find that county-level composition of the age of the mother, parity, and marital status were correlated with WIC exposure, but these differences are very small in magnitude. Having ruled out large amounts of compositional change in mothers with children, we turn to regressions stratifying on these 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 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 single and married mothers are not statistically different from one another by race). This evidence suggests that these individual risk factors are not likely to explain the differences in WIC impacts by race.

Fourth, in Figure 7, we explored place-based differences across counties at baseline. To do this, we divide the sample into those above and below the median for a given factor. To start with, we use the only county-level measure of take-up we could find. One of our Congressional reports had a list of WIC participating counties in 1980, along with take-up at the county level.[29] We predicted this take-up as a function of the baseline controls and then estimated effects for counties where this predicted take-up number was high or low. The effects are slightly larger at places with higher take-up, especially for Blacks.

Fifth, we explored additional place-based factors such as county-level poverty concentration, racial composition, and urbanicity in 1970, and found these factors to be of limited relevance. In addition, we explored the degree of medical need at the county level measured

[29]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.

by hospitals per capita and by the state share of Black participation in AFDC for women, but these results were also not very informative.

Lastly, we considered 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 as a measure of where hospitals were likely to be racially segregated. This set of states resembles the old Confederacy, and effects here seem to be larger in these places for Blacks but not for Whites. To summarize the different place-based factors we explored, we predicted the baseline average of low birth weight (in 1970) at the county level as a function of those county characteristics, with results not very conclusive.

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, therefore higher returns to investments for Black mothers relative to White mothers, and (2) racial differences in WIC program take-up and eligibility.

## 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.8. In Column 1, we show the results from our preferred specification. Next, we present specifications that only control for county and year-fixed effects (Column 2) and progressively add War on Poverty program covariates (Column 3), county time-varying socio-demographic 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.[30]

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 Figure

[30] Across different specifications, the point estimates on low birth weight range between 0.2 to 0.4 percentage points for Blacks and between 0.02 and 0.05 percentage points for Whites. Birth weight point estimates range between 5 to 13 grams for Blacks and between 2 to 5 grams for Whites.

A.3).[31] 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 the county composition may be different by race. To assess this issue, Appendix Figure A.3 presents estimates with a balanced set of counties in distant leads and lags, which are notably comparable to the full sample of counties.

Fourth, in Appendix Table A.9, we test whether our results are sensitive to weighting the estimates by the cell size (births in a cell). For this, we show our preferred specification, which uses propensity score reweighting and multiplying 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 and for a balanced set of counties with at least 2 Black and White births. As expected, the results for the all-county sample are sensitive to weighting by births because we effectively 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 number of births.

Fifth, we have performed additional robustness checks in Appendix Table A.10 that include: 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). Our results using individual-level data in Columns 1 and 6 are the same as the cell-level results weighted by births. This is reassuring and expected. In Columns 2, 3, 4 for Black births and 7, 8 and 9 for White births, we perform several robustness checks using alternative assignment 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 the missing gestational age to construct the gestational age for the full sample following Currie and Schwandt (2013). However, our results are not sensitive to using 30 days for the average days in a month, assigning 39 weeks for the missing gestational age, or keeping the gestational age as missing in our sample. Lastly, our results are very similar if we assign the WIC exposure in the month-year 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. In particular, we assign exposure based on the year of the second trimester, the year of the third trimester, and the year of birth (Appendix Table A.11). These estimates are highly comparable to those from our preferred specification.

Sixth, we check if our estimation is sensitive to trimming the propensity score to address

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

potential concerns about counties with a high propensity of having WIC affecting the results or controlling for the presence of the Commodity Supplemental Food Program (CSFP), which provided nutritional assistance to low-income families. We present these results in Appendix Table A.12. In Columns 1 and 4, we show our preferred specification, not trimming propensity score and controlling for CSFP presence. In Columns 2 and 5, we report that the results are not sensitive to trimming the propensity score by dropping counties with more than 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.

Seventh, our compilation of WIC rollout includes programs that served special populations only, including migrants and Native Americans. In Appendix Table A.13 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).

Lastly, since we build upon HPS WIC rollout study, we present estimations that replicate their specifications using their sample construction steps and years.[32] These results are presented in Appendix Table A.14. They use TWFE models and present two specifications. The first specification includes county and year fixed effects, time-varying per capita county transfer, per capita income and Food Stamps 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 estimates are similar to those reported in their paper. We corroborate their finding of an increase in birth weight but no effect on the probability of low birth weight. Columns 3 and 4 use their rollout but 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 county sample. 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 sample of counties and the point estimates for birth weight and low birth weight are larger.

## 6 Conclusion

TBD.

[32] 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 specification. The reason to exclude 1976 and 1977 years is the lack of primary sources of WIC rollout during those years. In addition, they limit their analysis to the set of counties that have WIC in place by 1979. WIC exposure is assigned in the year of the third trimester. The sample excludes cells that are less than 25 births 1971-1975 and 1978-1982.

## 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.

Bitler, M. P. and J. Currie (2005). Does wic work? the effects of wic on pregnancy and birth outcomes. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management 24*(1), 73–91.

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.

Butler, A. S. and R. E. Behrman (2007). *Preterm birth: causes, consequences, and prevention*. National academies press.

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.

Council, N. R., D. of Behavioral, S. Sciences, C. on Child Development Research, P. Policy, and P. for the Study of the Policy Formation Process (1982). *Making policies for children: A study of the federal process*. National Academies Press.

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, forthcoming*.

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.

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. (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. 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 (2022). Does the delivery of primary health care improve birth outcomes? Evidence from the rollout of community health services. Technical report, NBER WP 30047.

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.

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, Agricultureal 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.

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.

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

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.

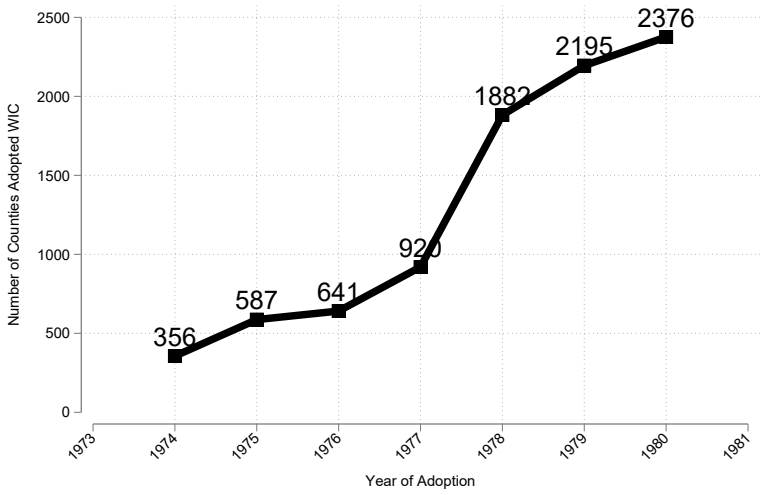
Singh, G. K. and S. M. Yu (2019). Infant Mortality in the United States, 1915-2017: Large Social Inequalities have Persisted for Over a Century. *International Journal of Maternal and Child Health and AIDS (IJMA) 8*(1), 19–31.

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

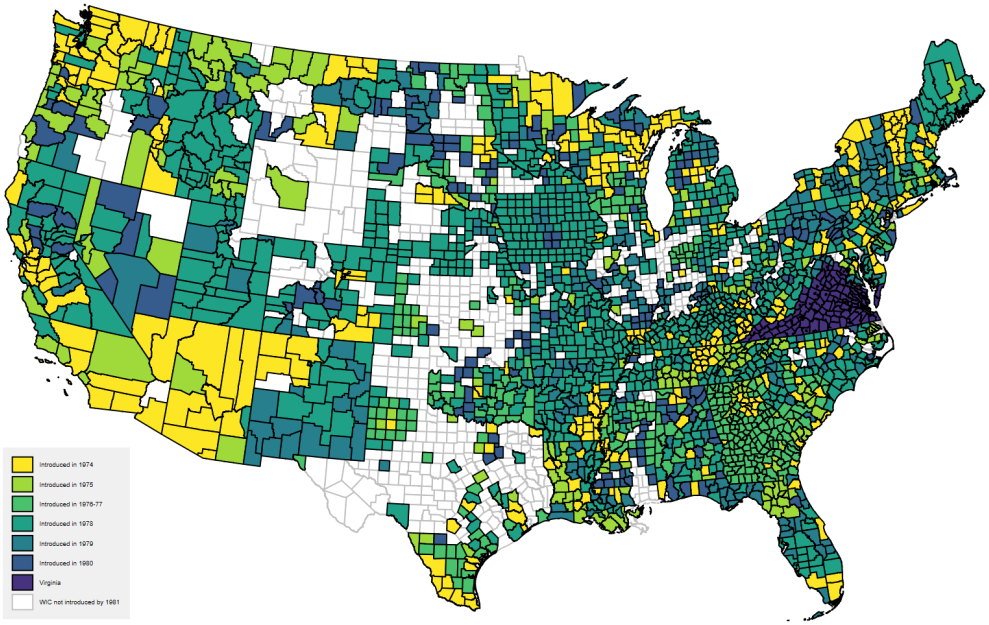
Vu, H., T. L. Green, and L. E. Swan (2024). Born on the wrong side of the tracks: Exploring the causal effects of segregation on infant health. *Journal of Health Economics 95*, 102876.

# Figures

**Figure 1:** Geographic Variation and Timing of WIC Adoption



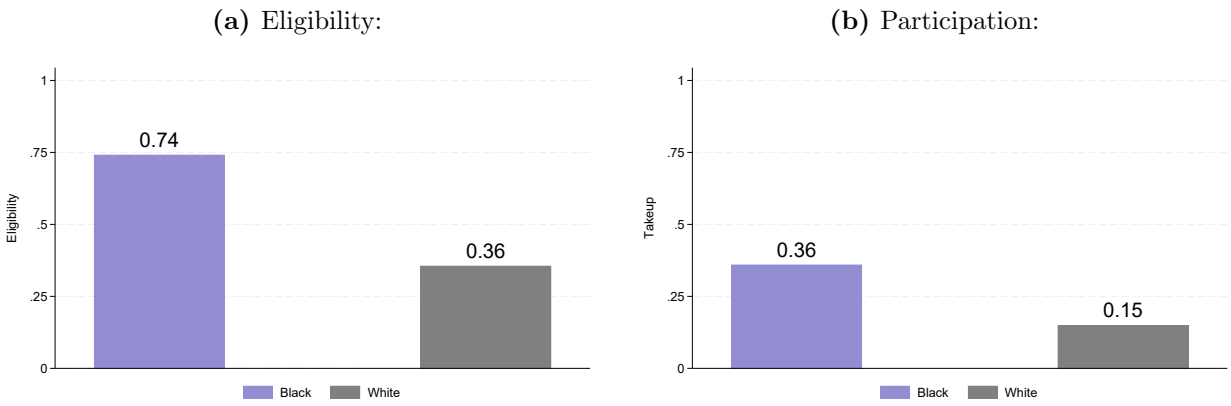

**(a)** Number of Counties Adopting WIC over Time




**(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, or Hawaii in our study.

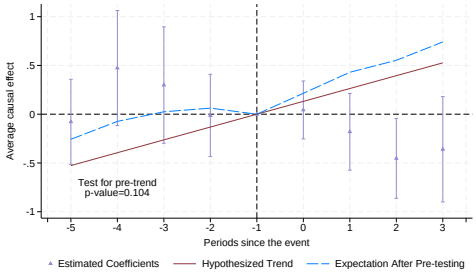
**Figure 2:** WIC Eligibility and Participation Are Higher among Blacks than Whites



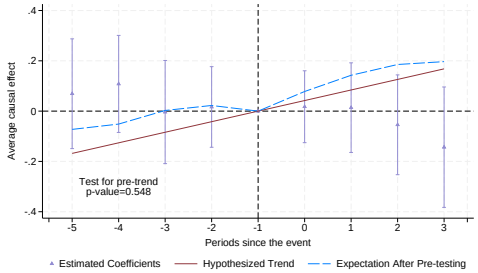

Notes: Eligibility calculations use measures of participation in WIC by race (Black, White, Asian/Pacific Islander, Native American, and Hispanic), and also broken down into women, infants, and children for Fiscal Year 1986 from (Food Research and Action Center, 1988). First, we combine the ITO counts 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 participants times the share of participants in the state (note that the share of lactating/postpartum participants is quite low across our time period), combining Hispanics with Whites for a "White" measure similar to 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 from the March CPS for 1986-1987 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 the previous year) with their CPS family (pooling related subfamilies with the heads) having total income under the upper-income limit at the time from (Food Research and Action Center, 1988) or being on AFDC or getting food stamps or on free or reduced school lunch or on Medicaid, summed with 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 report neither Medicaid use nor break out totals by race before that.

**Figure 3:** Event Study: The Effect of WIC on Main Birth Outcomes 1968–1981: Sun and Abraham (2021) Estimates, Propensity Score Specification - Black Mothers

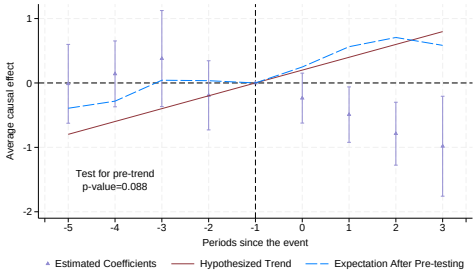
**(a)** Low Birth Weight



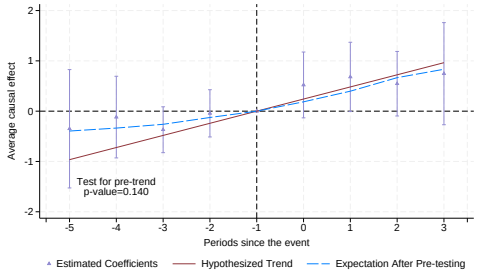

**(b)** Very Low Birth Weight



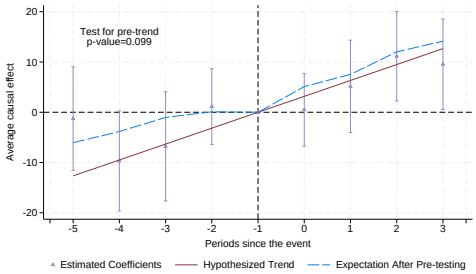

**(c)** Small for Gestational Age



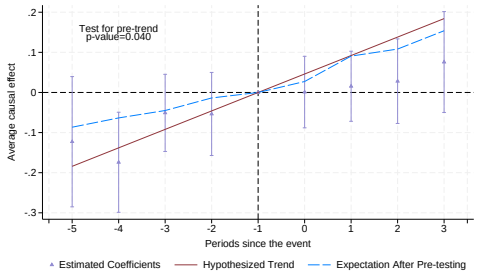

**(d)** Preterm



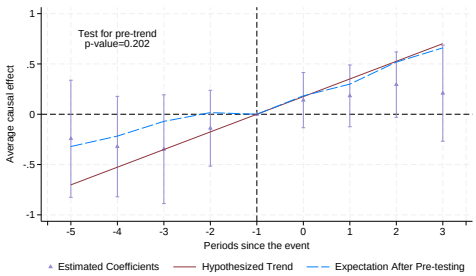

**(e)** Birth Weight



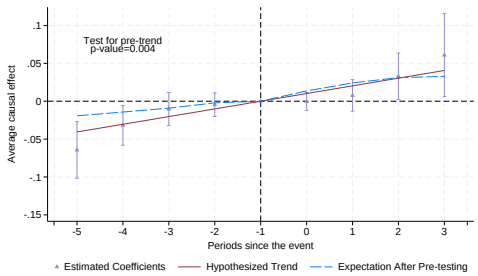

**(f)** High Birth Weight




**(g)** Large for Gestational Age



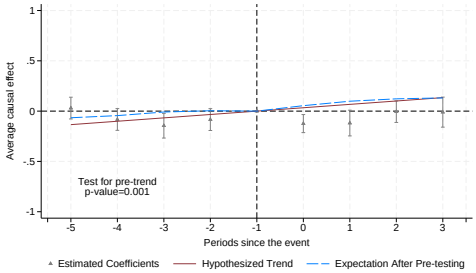

**(h)** Kessner Index-Adequate Prenatal Care



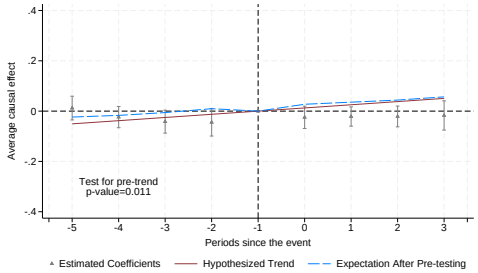

Notes: Figures report the estimated coefficients and standard errors using Sun and Abraham (2021) method. In addition, we plot potential violations of parallel trends based on methods Roth (2022) and report a p-value for a joint significant test for the pre-period coefficients. The red solid line represents a hypothesized linear violation of the parallel trends with 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 truth 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. The outcome data are from the Vital Statistical Natality Files that cover the years between 1968 and 1981. 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. * p<0.10, ** p<0.05, *** p<0.01.

**Figure 4:** Event Study: The Effect of WIC on Main Birth Outcomes 1968–1981: Sun and Abraham (2021) Estimates, Propensity Score Specification - White Mothers

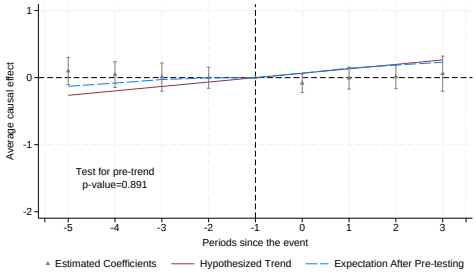
**(a)** Low Birth Weight



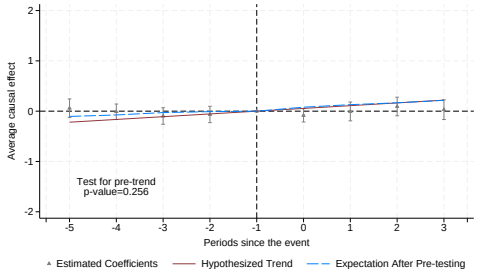

**(b)** Very Low Birth Weight



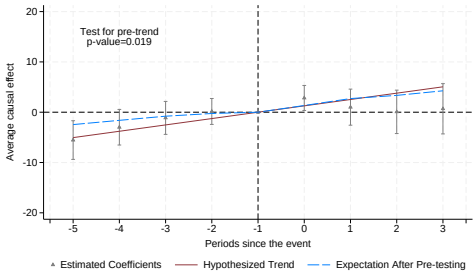

**(c)** Small for Gestational Age



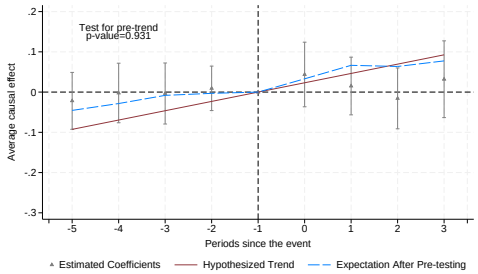

**(d)** Preterm



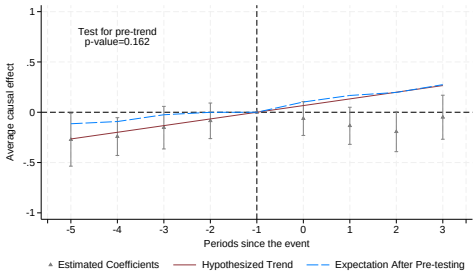

**(e)** Birth Weight



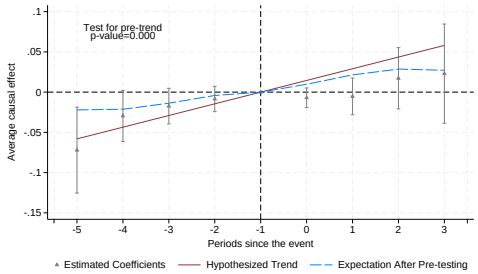

**(f)** High Birth Weight




**(g)** Large for Gestational Age



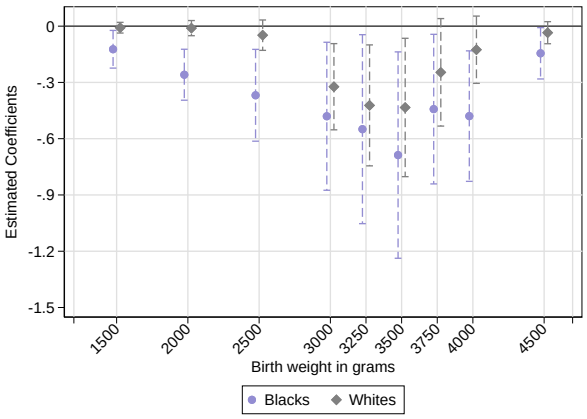

**(h)** Kessner Index-Adequate Prenatal Care



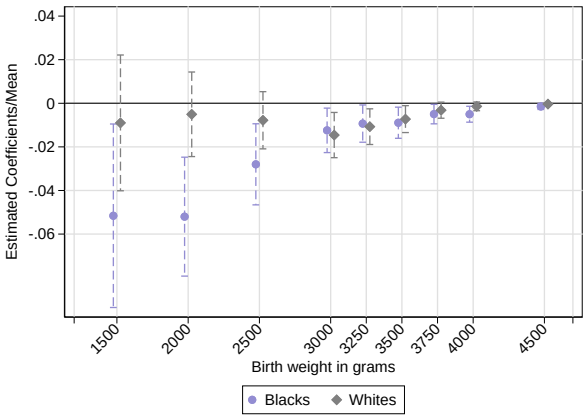

Notes: Figures report the estimated coefficients and standard errors using Sun and Abraham (2021) method. In addition, we plot potential violations of parallel trends based on methods Roth (2022) and report a p-value for a joint significant test for the pre-period coefficients. The red solid line represents a hypothesized linear violation of the parallel trends with 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 truth 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. The outcome data are from the Vital Statistical Natality Files that cover the years between 1968 and 1981. 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. * p<0.10, ** p<0.05, *** p<0.01.

**Figure 5:** Effects of WIC on Various Cut Points in the Distributional of Birth Weight

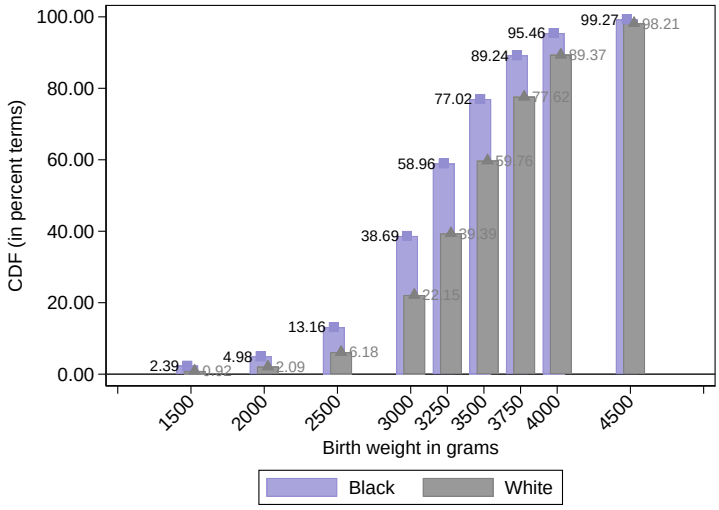
**(a)** Point Estimates




**(b)** Percent Impacts

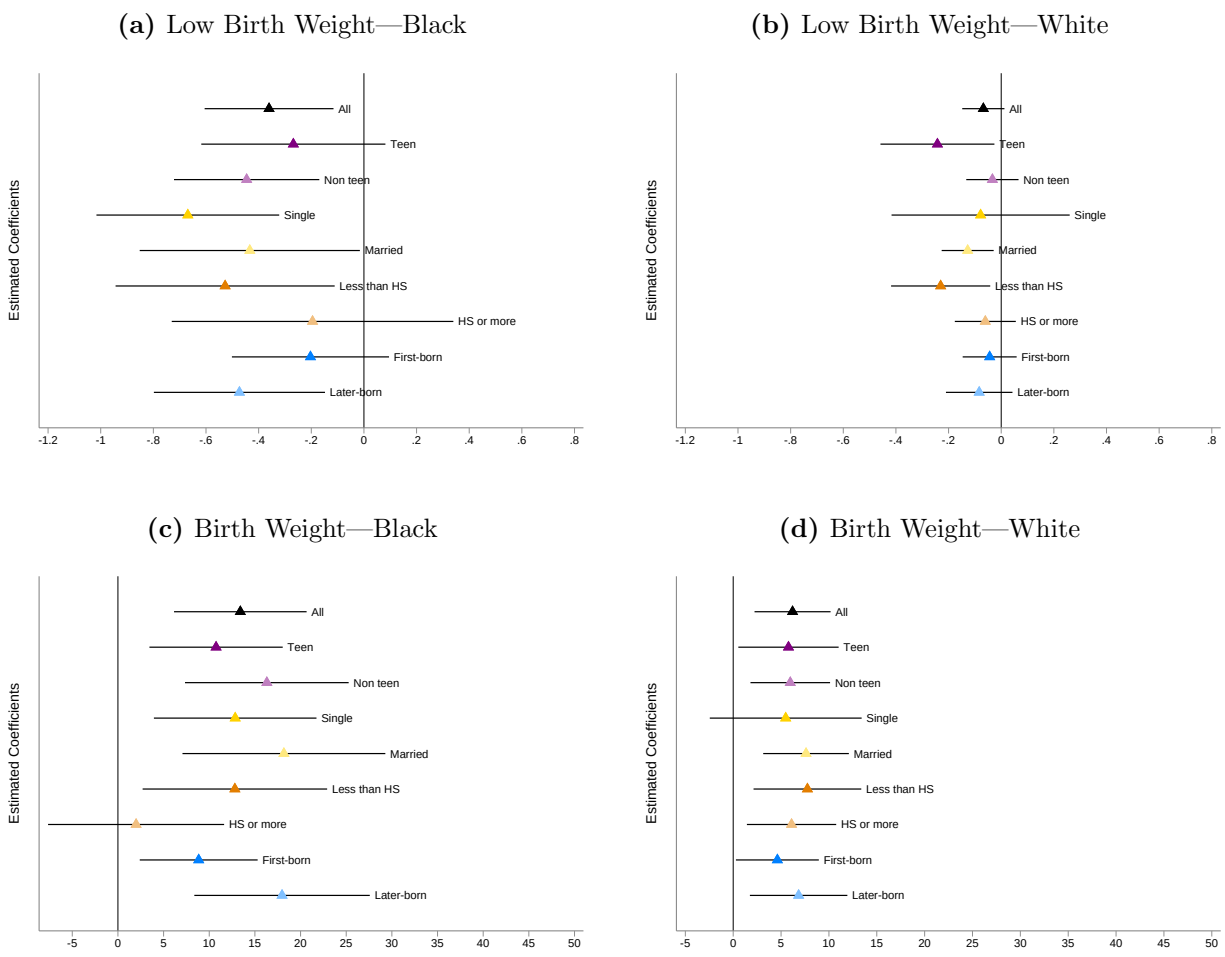



**(c)** Cumulative Distribution Function at the Birth Weight Cuts




Notes: We estimate the effect of WIC on the probability that birth weight is below a given gram threshold using individual level data: 1,500; 2,000; 2,500; 3,000; 3,250; 3,500; 3,750; 4,000; or, for 4,500 grams, the probability that it is above that number (Almond et al., 2011). We plot the point estimates and the 95% confidence intervals and "% Impacts [coef./mean]" separately using Sun and Abraham (2021) method and our preferred specification which re-weights untreated counties using an estimate of the propensity of WIC adoption based on pre-program characteristics.

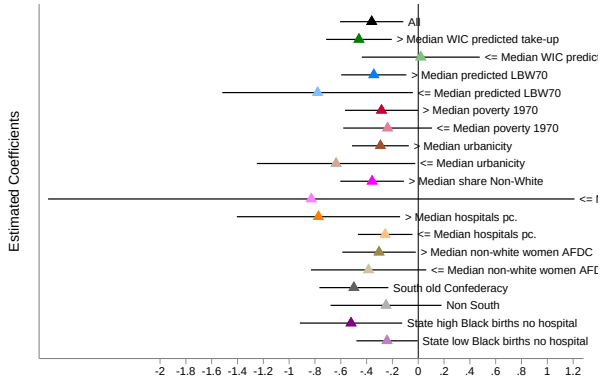
**Figure 6:** Heterogeneity Analyses: Individual Risk Factors



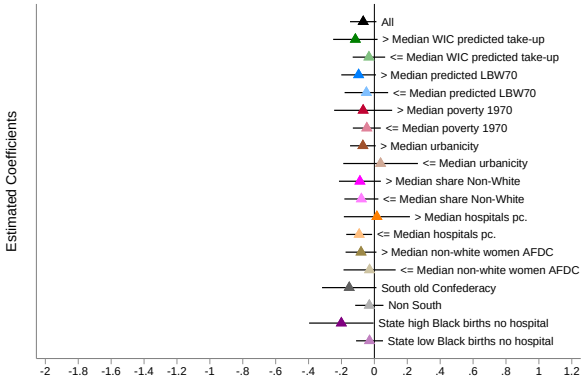

Notes: Static estimates using Sun and Abraham (2021) method and our preferred specification which re-weights untreated counties using an estimate of the propensity of WIC adoption based on pre-program characteristics. Each row presents an estimate for a single subsample (triangle) along with 95% confidence levels, for Blacks (left) and Whites (right), for low birth weight (upper panels) and birth weight (lower panels). Each row is from a separate regression.

**Figure 7:** Heterogeneity Analyses—Place-Based Factors

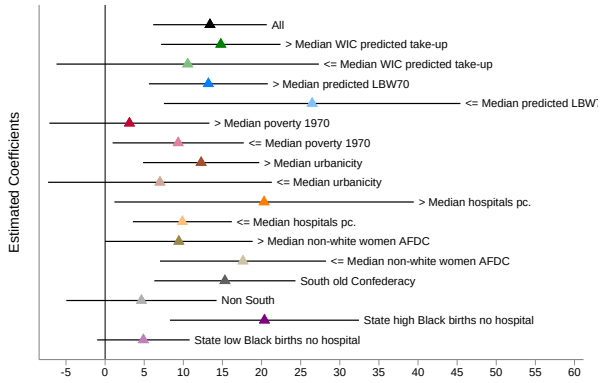
**(a)** Low Birth Weight—Black



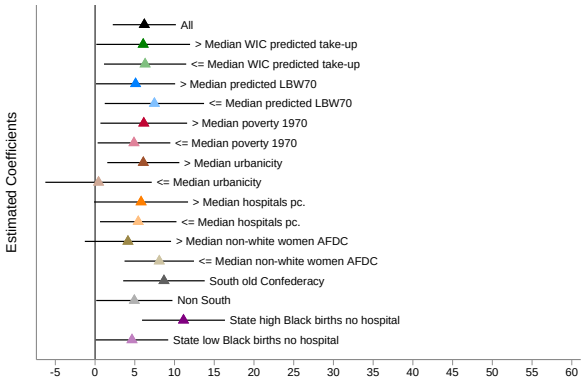

**(b)** Low Birth Weight—White




**(c)** Birth Weight—Black




**(d)** Birth Weight—White




Notes: Static estimates using Sun and Abraham (2021) method and our preferred specification which re-weights untreated counties using an estimate of the propensity of WIC adoption based on pre-program characteristics. Each row presents an estimate for a single subsample (triangle) along with 95% confidence levels, for Blacks (left) and Whites (right), for low birth weight (upper panels) and birth weight (lower panels). Each row is from a separate regression.

# Tables

**Table 1:** The Effect of WIC on Main Birth Outcomes 1968–1981 by Race-County-Year of Second Trimester Cells: Sun and Abraham (2021) Estimates, Propensity Score Specification

| | 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 | -0.3685*** | -0.1232** | -0.7234** | 0.9084* | 12.7268*** | 0.1436** | 0.6081*** | 0.0240* |
| | (0.1249) | (0.0513) | (0.3074) | (0.5432) | (3.8047) | (0.0700) | (0.2292) | (0.0140) |
| Control Y-mean | 13.431 | 2.361 | 14.455 | 17.874 | 3080.177 | 0.733 | 6.049 | 0.189 |
| % Effect (Coef/Control Mean) | -2.744 | -5.217 | -5.004 | 5.082 | 0.413 | 19.587 | 10.053 | 12.684 |
| Adj. R-Squared | 0.025 | 0.013 | 0.039 | 0.058 | 0.072 | 0.028 | 0.040 | 0.626 |
| Obs | 205159 | 205159 | 158481 | 158481 | 205159 | 205159 | 158481 | 163528 |
| *B: White Mothers* | | | | | | | | |
| WIC-month year-trim2 | -0.0481 | -0.0082 | -0.0831 | 0.0471 | 4.8842** | 0.0310 | 0.1530* | 0.0048 |
| | (0.0414) | (0.0146) | (0.1034) | (0.0839) | (2.0338) | (0.0303) | (0.0861) | (0.0206) |
| Control Y-mean | 6.425 | 0.932 | 8.610 | 8.424 | 3339.554 | 1.667 | 10.094 | 0.288 |
| % Effect (Coef/Control Mean) | -0.749 | -0.883 | -0.965 | 0.559 | 0.146 | 1.857 | 1.516 | 1.682 |
| Adj. R-Squared | 0.051 | 0.007 | 0.065 | 0.042 | 0.264 | 0.046 | 0.073 | 0.706 |
| Obs | 474383 | 474383 | 412460 | 412460 | 474383 | 474383 | 412460 | 380766 |
| 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 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 untreated cells are weighted by the number of births in the cell (treatment group) and $\hat{p}/(1-\hat{p})$ times the number of births in the cell (control group). The outcome data are from Vital Statistical Natality Files that cover the years between 1968 and 1981. 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. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

**Table 2:** The Effect of WIC on Composition of Births 1968–1981 by Race-County-Year of Second Trimester Cells: Sun and Abraham (2021) Estimates, Propensity Score Specification

| | 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 | 0.0072 | -0.0025 | 0.0206** | 0.0616** | 0.1494 | -1.2130* | 0.6962 | -0.5336 | -0.1627 |
| | (0.0497) | (0.0164) | (0.0094) | (0.0276) | (0.2123) | (0.6276) | (0.4535) | (0.4846) | (0.2784) |
| Y-mean | 9.106 | 5.562 | 2.101 | 24.392 | 18.013 | 87.697 | 27.708 | 46.874 | 25.418 |
| % Effect (Coef/Control Mean) | 0.079 | -0.045 | 0.981 | 0.253 | 0.829 | -1.383 | 2.513 | -1.138 | -0.640 |
| Adj. R-Squared | 0.834 | 0.983 | 0.449 | 0.610 | 0.482 | 0.784 | 0.739 | 0.458 | 0.729 |
| Obs | 475320 | 475320 | 476328 | 476328 | 476328 | 334490 | 318285 | 318285 | 318285 |
| 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 Sun and Abraham (2021) method, controlling for county and year of the second trimester. The outcome data are from the Vital Statistical Natality Files that cover the years between 1968 and 1981. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level, 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. The outcome data are from the Vital Statistical Natality Files that cover the years between 1968 and 1981. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level collapsed at county and year of the second trimester. Standard errors are clustered at the county level. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

**Table 3:** Robustness to Specification: The Effect of WIC on Low Birth Weight 1968–1981 by Race-County-Month and Year of Second Trimester Cells: Sun and Abraham (2021) and TWFE Estimates, All Counties

| | Sun and Abraham (2021) | | | | | | TWFE |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| *A: Black Mothers* | | | | | | | |
| WIC-month year-trim2 | -0.3685*** | -0.2920*** | -0.2955*** | -0.2764*** | -0.2224** | -0.2461*** | -0.3500*** |
| | (0.1249) | (0.0862) | (0.0825) | (0.0828) | (0.0874) | (0.0845) | (0.1185) |
| Control Y-mean | 13 | 13 | 13 | 13 | 13 | 13 | 13 |
| Adj. R-Squared | 0.025 | 0.028 | 0.028 | 0.028 | 0.028 | 0.030 | 0.025 |
| Obs | 205159 | 205159 | 205159 | 205123 | 205123 | 205123 | 205159 |
| N States | 48 | 48 | 48 | 48 | 48 | 48 | 48 |
| N Counties | 2190 | 2190 | 2190 | 2190 | 2190 | 2190 | 2190 |
| Min Cell Size | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Max Cell Size | 3046 | 3046 | 3046 | 3046 | 3046 | 3046 | 3046 |
| *B: White Mothers* | | | | | | | |
| WIC-month year-trim2 | -0.0481 | -0.0330 | -0.0308 | -0.0310 | -0.0163 | -0.0187 | -0.0406 |
| | (0.0414) | (0.0224) | (0.0224) | (0.0225) | (0.0235) | (0.0232) | (0.0407) |
| Control Y-mean | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| Adj. R-Squared | 0.051 | 0.055 | 0.055 | 0.055 | 0.056 | 0.057 | 0.051 |
| Obs | 474383 | 474383 | 474383 | 474335 | 474335 | 474335 | 474383 |
| N States | 48 | 48 | 48 | 48 | 48 | 48 | 48 |
| N Counties | 2886 | 2886 | 2886 | 2886 | 2886 | 2886 | 2886 |
| Min Cell Size | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Max Cell Size | 9671 | 9671 | 9671 | 9671 | 9671 | 9671 | 9671 |
| 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 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 untreated cells are weighted by the number of births in the cell (treatment group) and $\hat{p}/(1-\hat{p})$ times the number of births in the cell (control group). The outcome data are from Vital Statistical Natality Files that cover the years between 1968 and 1981. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level. 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.

# 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) for years between 1978 and 1980 for some states not included in the HPS data.
- 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 planes, data sources used to digitize WIC implementation included:
  - 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 validate 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, "internal conflict"
    * If at least two out of three sources of information indicated the presence of WIC, the county in question had 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; 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 mostly run by Indian Tribal Organizations 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 boundary was served by WIC

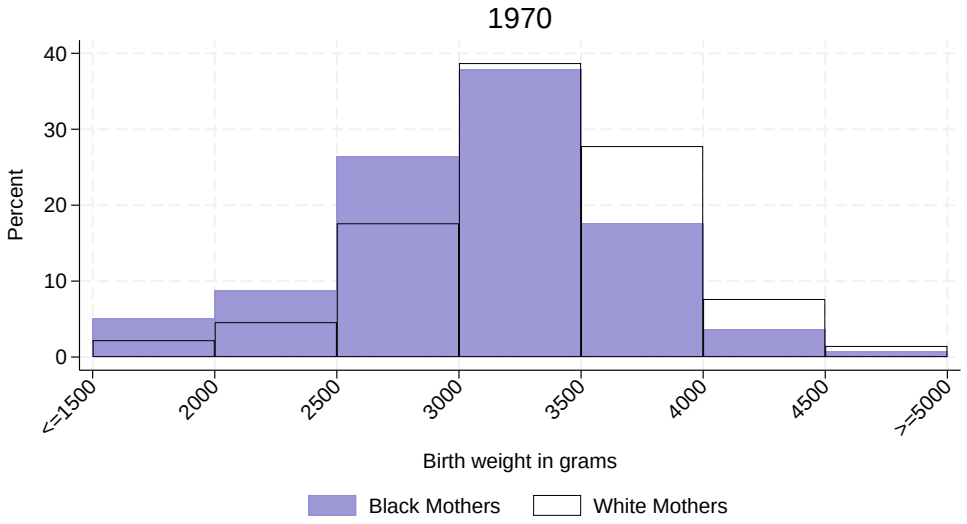
- A detailed explanation of the digitization process by fiscal year and state including raw source information for counties with conflicting WIC indications was maintained by the research team; this includes all internal conflict counties, special population counties, and decisions regarding the presence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

## B Sample Restrictions

We begin by using all the observations in the Vital Statistics Natality Data. We exclude births 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. We include some results for the births to non-Native American, non-Black, and Non-White mothers in the Appendix tables. The results for Other mothers tend to be between those for Blacks and Whites (Appendix Table A.7). Earlier states and years of the data are a 50% sample of birth certificates, so we weight these observations by 2. We exclude twin and higher order births in the years this is available and also exclude mothers who do not reside in a US state. We code according to WIC availability in 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, some of the components of adequacy of prenatal care are not available for all states and years of occurrence.

# Appendix Figures

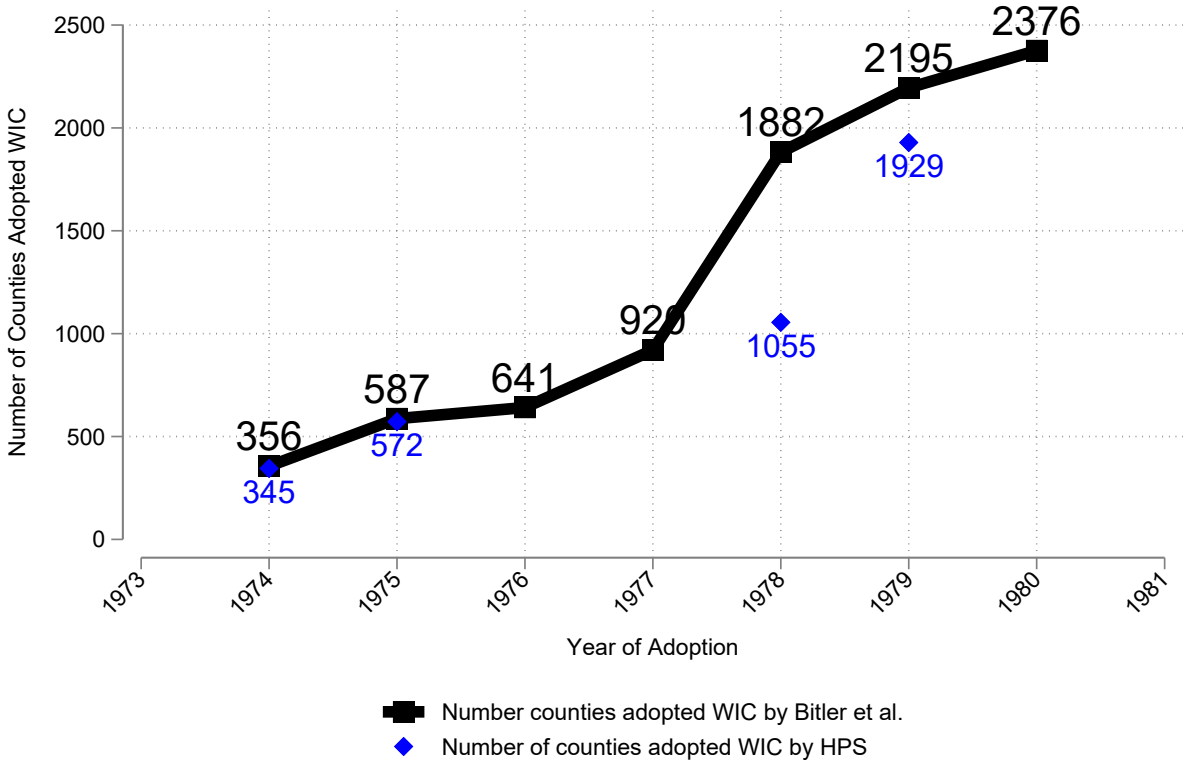
**Figure A.1:** Birth Weight Distribution in 1970




Source: Vital Statistics Birth Records

**Figure A.2:** Number of Counties Adopting WIC over Time, between 1974 and 1980:

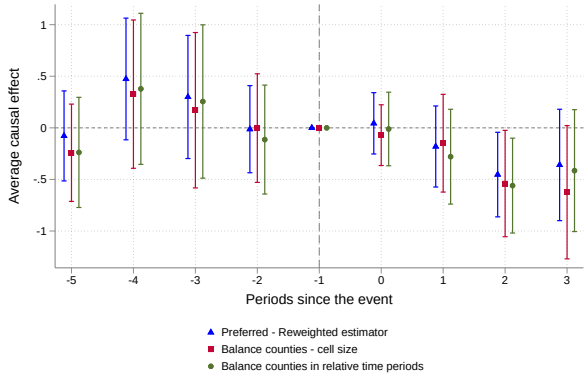
Our Data Compilation vs. Hoynes et al. (2011)



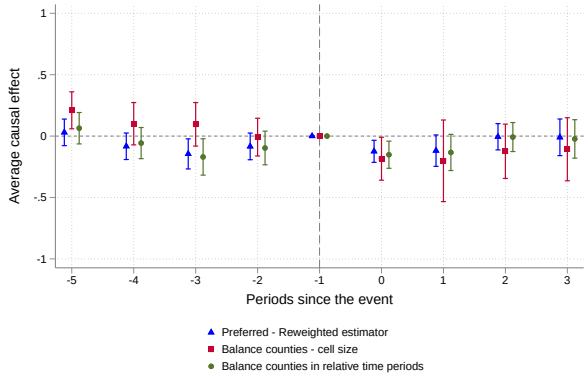

Notes: Hoynes et al. (2011) (HPS) generously shared their data for the counties in their sample. HPS had information for 1978 for 13 states' migrant programs only and for 1979 information from a WIC directory. The State plans include detailed information on local agencies and counties served and add WIC roll-out information for the years 1978, 1979, and 1980.

**Figure A.3:** Event Study: Robustness Checks using Propensity Score Weights for Control Counterfactual and Alternative Samples

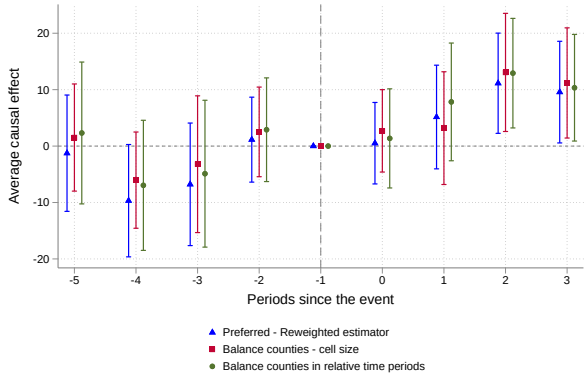
**(a)** Black Mothers: Low Birth Weight



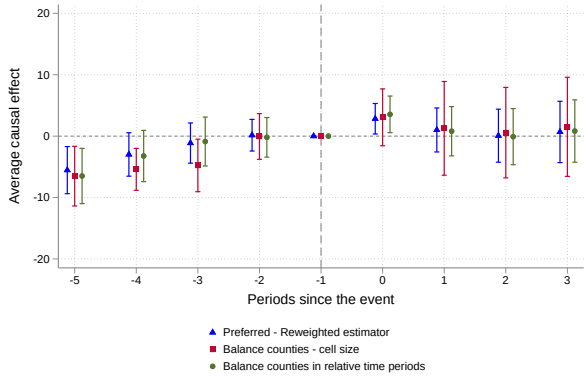

**(b)** White Mothers: Low Birth Weight




**(c)** Black Mothers: Birth Weight



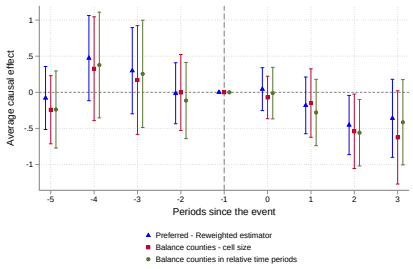

**(d)** White Mothers: Birth Weight



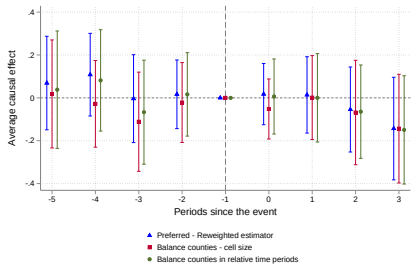

Notes: Figures report the estimated coefficients and standard errors using the Sun and Abraham (2021) method. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption, $\hat{p}$, by 1980, based on pre-program county characteristics. The cells are weighted by the number of births in the cell (treatment group) and $\hat{p}/(1-\hat{p})$ times the number of births in the cell (control group). The outcome data are from the Vital Statistical Natality Files that cover the years between 1968 and 1981. We compile data on WIC timing from multiple sources between 1974 and 1980. Standard errors are clustered at the county level. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

**Figure A.4:** Event Study: The Effect of WIC on Main Birth Outcomes 1968–1981: Sun and Abraham (2021) Estimates, Propensity Score Specification - Black Mothers

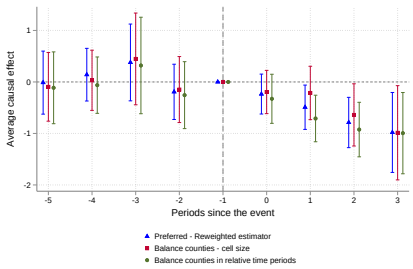
**(a)** Low Birth Weight



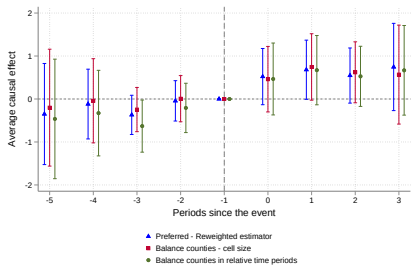

**(b)** Very Low Birth Weight



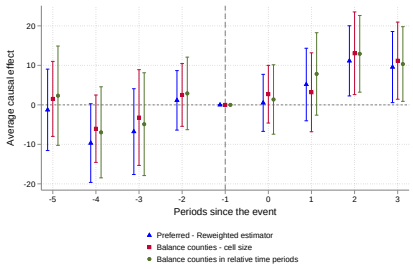

**(c)** Small for Gestational Age



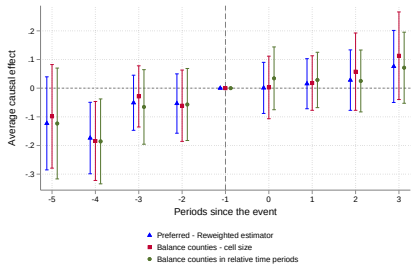

**(d)** Preterm



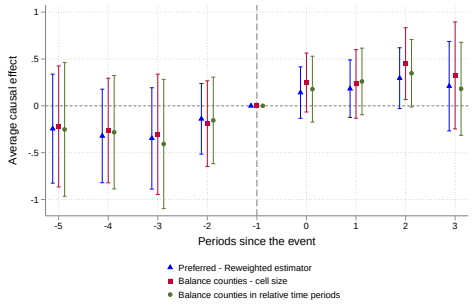

**(e)** Birth Weight



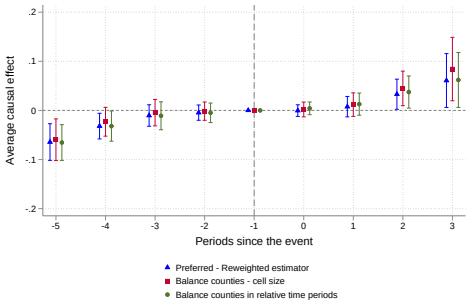

**(f)** High Birth Weight




**(g)** Large for Gestational Age



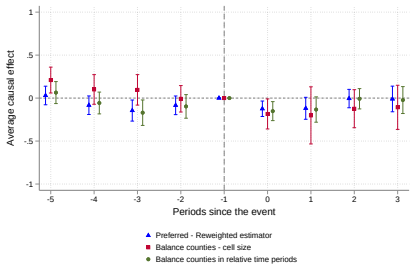

**(h)** Kessner Index Adequate



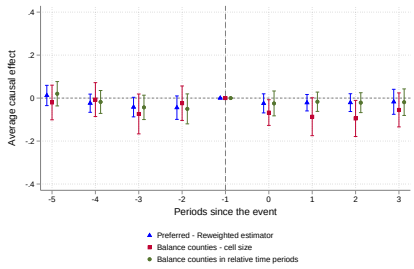

Notes: Figures report the estimated coefficients and standard errors using Sun and Abraham (2021) method. Our preferred specification re-weights untreated counties using an estimate of the propensity of WIC adoption based on pre-program characteristics. The outcome data are from the Vital Statistical Natality Files that cover the years between 1968 and 1981. 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. * p<0.10, ** p<0.05, *** p<0.01.

**Figure A.5:** Event Study: The Effect of WIC on Main Birth Outcomes 1968–1981: Sun and Abraham (2021) Estimates, Propensity Score Specification - White Mothers

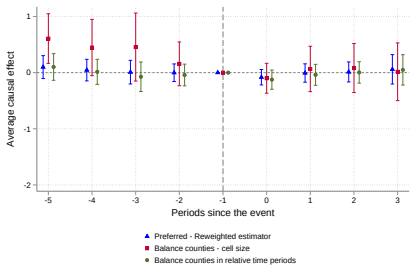
**(a)** Low Birth Weight



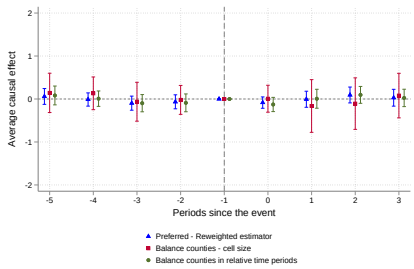

**(b)** Very Low Birth Weight



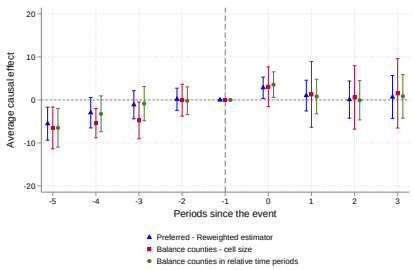

**(c)** Small for Gestational Age



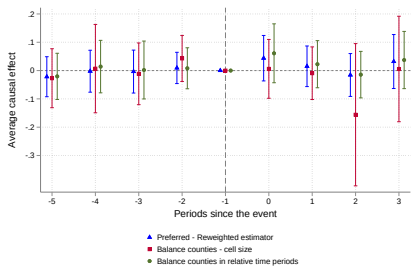

**(d)** Preterm



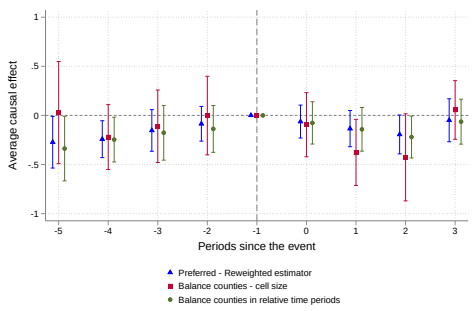

**(e)** Birth Weight



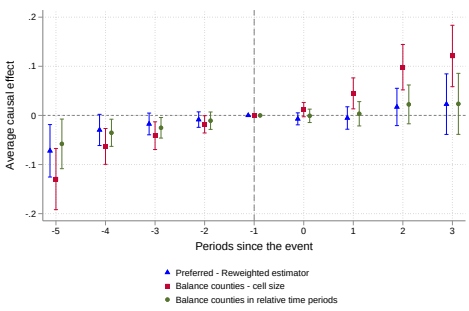

**(f)** High Birth Weight




**(g)** Large for Gestational Age

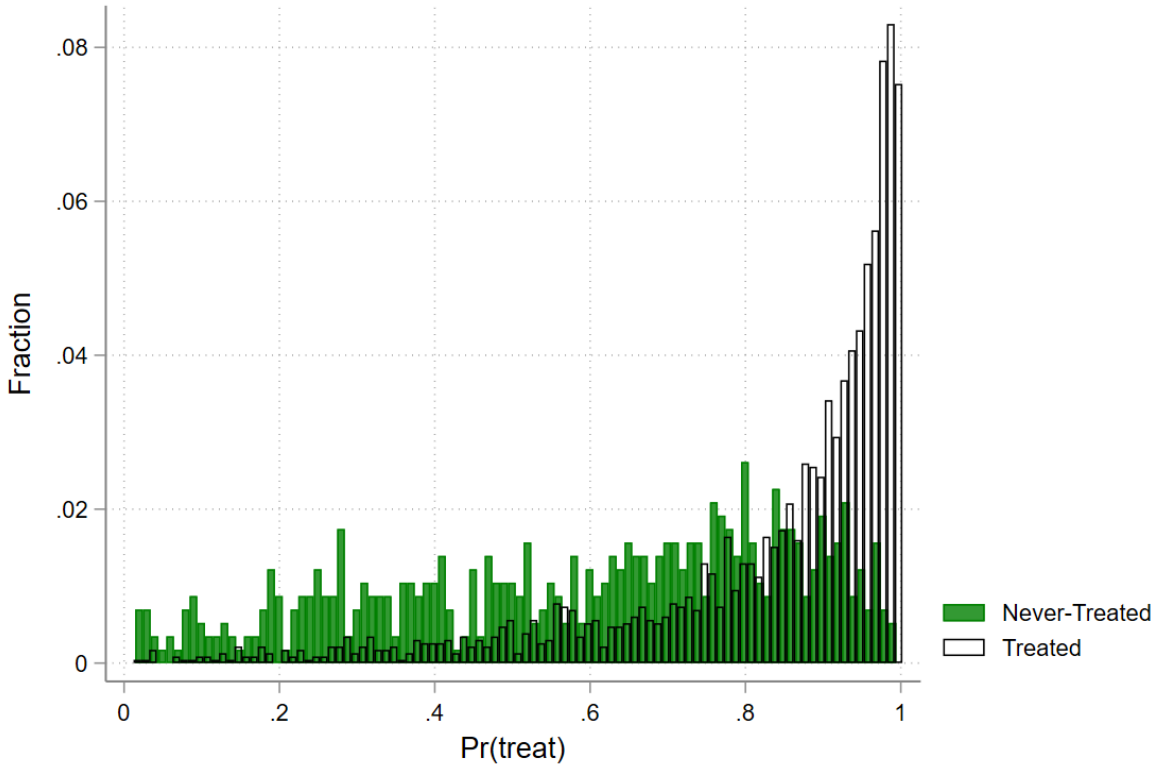



**(h)** Kessner Index Adequate




Notes: Figures report the estimated coefficients and standard errors using 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. The outcome data are from the Vital Statistical Natality Files that cover the years between 1968 and 1981. 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. * p<0.10, ** p<0.05, *** p<0.01.

**Figure A.6:** Common Support for the Propensity Score




Notes: Figure reports the distribution of the predicted propensity to adopt WIC using pre-program county characteristics as predictors.

# Appendix Tables

**Table A.1:** Cross-Tabulation of WIC Roll-out: Bitler et al. vs. Hoynes et al. (2011)

| | WIC Roll-out HPS | | | | | |
|---|---|---|---|---|---|---|
| WIC Roll-out Our Compilation | 1974 | 1975 | 1978 | 1979 | _missing_ | 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_ | 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 Roll-out 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. Of note is that HPS rollout years exclude 1977 and 1976 given their lack of primary sources for that period. The sample corresponds to WIC counties in HPS or in our rollout data excluding Virginia and Hawaii.

**Table A.2:** Summary Statistics for Outcomes and Controls for Full Sample and by WIC Adoption Status using Vital Statistics Data 1968–1981, Collapsed by Race-County-Month and Year of Second Trimester Cells

| | All Counties | | Counties with WIC | | Counties without WIC | |
|---|---|---|---|---|---|---|
| | Black | White | Black | White | Black | White |
| ***Outcomes*** | | | | | | |
| Low Birth Weight | 13.161 | 6.176 | 13.147 | 6.172 | 13.367 | 6.226 |
| | (7.726) | (3.630) | (7.470) | (3.406) | (10.775) | (5.683) |
| Very Low Birth Weight | 2.388 | 0.916 | 2.398 | 0.919 | 2.244 | 0.882 |
| | (3.384) | (1.390) | (3.277) | (1.309) | (4.662) | (2.135) |
| Small for Gest. Age | 14.801 | 8.496 | 14.782 | 8.488 | 15.091 | 8.588 |
| | (10.834) | (7.941) | (10.832) | (7.924) | (10.858) | (8.159) |
| Preterm | 17.267 | 8.163 | 17.257 | 8.196 | 17.423 | 7.732 |
| | (13.563) | (9.715) | (13.655) | (9.825) | (12.099) | (8.153) |
| Birth Weight | 3088.061 | 3355.619 | 3088.137 | 3355.520 | 3086.968 | 3356.826 |
| | (147.408) | (99.967) | (142.821) | (94.710) | (202.568) | (149.718) |
| High Birth Weight | 0.732 | 1.787 | 0.726 | 1.784 | 0.815 | 1.817 |
| | (2.106) | (1.985) | (2.022) | (1.864) | (3.077) | (3.097) |
| Large for Gest. Age | 5.866 | 10.230 | 5.866 | 10.219 | 5.877 | 10.364 |
| | (7.377) | (8.175) | (7.363) | (8.140) | (7.585) | (8.625) |
| Kessner Index - Adequate | 0.278 | 0.379 | 0.279 | 0.374 | 0.273 | 0.442 |
| | (0.247) | (0.334) | (0.248) | (0.335) | (0.234) | (0.324) |
| ***WOP exposure*** | | | | | | |
| Food Stamps Exposure | 0.908 | 0.873 | 0.910 | 0.876 | 0.883 | 0.844 |
| | (0.289) | (0.333) | (0.287) | (0.330) | (0.322) | (0.363) |
| CHC Exposure | 0.497 | 0.318 | 0.502 | 0.337 | 0.423 | 0.095 |
| | (0.500) | (0.466) | (0.500) | (0.473) | (0.494) | (0.293) |
| ***Expenditures, Hospitals and Population*** | | | | | | |
| UI State Expenditures per capita | 126.574 | 133.683 | 130.029 | 137.920 | 76.483 | 82.265 |
| | (109.269) | (109.223) | (111.307) | (110.106) | (52.332) | (81.964) |
| Medicare Expenditures per capita | 283.718 | 270.373 | 285.301 | 271.684 | 260.788 | 254.465 |
| | (144.530) | (125.634) | (146.542) | (126.030) | (108.842) | (119.586) |
| Number of Hospitals per capita | 0.029 | 0.032 | 0.028 | 0.030 | 0.040 | 0.052 |
| | (0.021) | (0.031) | (0.019) | (0.027) | (0.034) | (0.058) |
| Population (000s) | 1160.639 | 785.332 | 1216.002 | 839.957 | 357.830 | 122.264 |
| | (1791.923) | (1545.130) | (1838.064) | (1594.331) | (305.466) | (191.192) |
| ***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 | 13.793 | 10.614 | 13.441 | 10.525 | 18.906 | 11.698 |
| | (9.114) | (6.818) | (9.177) | (6.837) | (6.194) | (6.490) |
| N | 205159 | 474383 | 179028 | 381445 | 26131 | 92938 |

Notes: The outcomes are from the Vital Statistics Natality Data from 1968–1981. The WOP exposure to Food Stamps and Community Health Centers comes from Hoynes et al. (2011) and Kose et al. (2022). The other controls come from Almond et al. (2011).

**Table A.3:** Determinants of WIC Roll-out Timing by County, Using 1970 County Characteristics

| | Unweighted | | Weighted | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| % population age <5-1970 | 0.020 | 0.035 | -0.117 | -0.209 |
| | (0.044) | (0.047) | (0.138) | (0.143) |
| % population aged 65+ -1970 | 0.078*** | 0.067*** | 0.020 | -0.035 |
| | (0.020) | (0.021) | (0.053) | (0.059) |
| Log 1970 population-1970 | -0.196*** | -0.362*** | -0.476*** | -0.484*** |
| | (0.053) | (0.057) | (0.106) | (0.118) |
| % below poverty level-1969 | -0.016* | -0.028*** | -0.040* | -0.047* |
| | (0.009) | (0.010) | (0.022) | (0.024) |
| % persons 25+ w/ <4 yrs sch-1970 | -0.041*** | -0.005 | -0.022 | 0.004 |
| | (0.015) | (0.016) | (0.029) | (0.030) |
| % persons 25+ w/ 12+ yrs sch-1970 | -0.000 | 0.015** | 0.009 | 0.032** |
| | (0.006) | (0.007) | (0.011) | (0.015) |
| % nonwhite-1970 | 0.001 | -0.000 | 0.003 | 0.002 |
| | (0.003) | (0.004) | (0.007) | (0.009) |
| % urban-1970 | -0.001 | 0.000 | -0.012** | -0.017*** |
| | (0.002) | (0.002) | (0.005) | (0.005) |
| % rural farm-1970 | -0.234 | 0.086 | -1.377* | -2.524** |
| | (0.381) | (0.468) | (0.757) | (1.166) |
| Medicare benefits pc (1000s 2006)-1970 | -0.001 | -0.001 | 0.001 | 0.001 |
| | (0.001) | (0.001) | (0.002) | (0.003) |
| Pub assist medical care benefits pc (1000s 2006)-1970 | -0.001** | -0.000 | -0.000 | -0.000 |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| Supplementary security income pc (1000s 2006)-1970 | 0.001 | -0.001 | 0.000 | 0.002 |
| | (0.001) | (0.001) | (0.002) | (0.003) |
| State UI compensation pc (1000s 2006)-1970 | -0.002*** | -0.001 | -0.001 | 0.001 |
| | (0.001) | (0.001) | (0.001) | (0.002) |
| Active MDs pc -1960 | -0.192** | -0.222*** | -0.320** | -0.299** |
| | (0.075) | (0.072) | (0.152) | (0.129) |
| Beds pc-1970 | 0.008 | 0.006 | 0.054** | 0.029 |
| | (0.010) | (0.009) | (0.024) | (0.023) |
| Hospitals pc-1970 | -0.002** | -0.001 | -0.006*** | -0.003 |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| Infant mortality rate-1970 | 0.002 | -0.000 | 0.009 | 0.009 |
| | (0.003) | (0.003) | (0.008) | (0.007) |
| Adopted FS before 1970 | -0.048 | -0.021 | 0.233 | 0.170 |
| | (0.074) | (0.094) | (0.159) | (0.185) |
| Adopted CHC before 1975 | -1.115*** | -1.036*** | -0.892*** | -0.737*** |
| | (0.204) | (0.211) | (0.280) | (0.266) |
| Adj. R-Squared | 0.111 | 0.223 | 0.351 | 0.423 |
| Obs | 2315 | 2315 | 2314 | 2314 |
| Weighted | N | N | Y | Y |
| State FE | | | | |

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. 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.4:** Balance Test by WIC Adoption Status across Counties: 1970 County Characteristics Summary Statistics

| | Raw | | | Weighted by pscore hat | | | Add state FE | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean Diff (1) | SE (2) | MHT p-val (3) | Mean Diff (4) | SE (5) | MHT p-val (6) | Mean Diff (7) | SE (clustered) (8) | MHT p-val (9) |
| % population age <5-1970 | 0.73*** | 0.07 | 0.01 | 0.07 | 0.09 | 0.99 | 0.17 | 0.12 | 0.98 |
| % population aged 65+ -1970 | -0.70*** | 0.07 | 0.01 | -0.11 | 0.08 | 0.90 | -0.08 | 0.16 | 0.98 |
| Log 1970 population-1970 | 0.45*** | 0.03 | 0.01 | 0.12** | 0.06 | 0.72 | 0.05 | 0.08 | 0.98 |
| % below poverty level-1969 | 0.37*** | 0.05 | 0.01 | 0.09 | 0.18 | 0.99 | 0.23 | 0.15 | 0.98 |
| % persons 25+ w/ <4 yrs sch-1970 | 0.54*** | 0.05 | 0.01 | -0.18 | 0.31 | 0.99 | 0.07 | 0.15 | 0.99 |
| % persons 25+ w/ 12+ ysch-1970 | -0.30*** | 0.05 | 0.01 | -0.07 | 0.11 | 0.99 | -0.16 | 0.12 | 0.98 |
| % nonwhite-1970 | 0.43*** | 0.04 | 0.01 | 0.20* | 0.11 | 0.73 | 0.10 | 0.13 | 0.98 |
| % urban-1970 | 0.21*** | 0.05 | 0.01 | 0.05 | 0.08 | 0.99 | 0.10 | 0.13 | 0.98 |
| % rural farm-1970 | -0.87*** | 0.09 | 0.01 | -0.22 | 0.17 | 0.91 | -0.18 | 0.13 | 0.98 |
| Medicare pc(1000s 2006)-1970 | -0.56*** | 0.06 | 0.01 | -0.17** | 0.08 | 0.66 | -0.06 | 0.10 | 0.98 |
| State UI pc (1000s 2006)-1970 | 0.55*** | 0.03 | 0.01 | 0.10 | 0.11 | 0.97 | -0.01 | 0.09 | 1.00 |
| Hospitals pc-1970 | -1.35*** | 0.16 | 0.01 | -0.38* | 0.20 | 0.74 | -0.25 | 0.18 | 0.98 |
| Active MDs pc -1960 | 0.09*** | 0.02 | 0.01 | 0.07** | 0.04 | 0.69 | 0.06 | 0.05 | 0.98 |
| % births not hospital -1970 | 0.23*** | 0.06 | 0.01 | -0.51 | 0.47 | 0.96 | -0.15 | 0.27 | 0.98 |
| Women AFDC rate 67-73 - White | 0.36*** | 0.03 | 0.01 | 0.23*** | 0.08 | 0.33 | 0.00 | | |
| Women AFDC rate 67-73 - Nonwhite | 0.29*** | 0.05 | 0.01 | 0.09 | 0.13 | 0.99 | 0.00 | | |
| Children AFDC rate 67-73-White | 0.33*** | 0.03 | 0.01 | 0.22*** | 0.07 | 0.28 | 0.00 | | |
| ChildrenAFDC rate 67-73-Nonwhite | 0.27*** | 0.05 | 0.01 | 0.12 | 0.10 | 0.94 | 0.00 | | |
| FS opened pre-1974 | -0.11** | 0.05 | 0.09 | 0.11 | 0.11 | 0.96 | 0.12 | 0.10 | 0.98 |
| CHC opened pre-1974 | 0.06*** | 0.01 | 0.01 | 0.07*** | 0.01 | 0.08 | 0.04 | 0.03 | 0.98 |
| CSFP pre-1974 | 0.07*** | 0.02 | 0.02 | 0.07* | 0.04 | 0.75 | 0.00 | 0.06 | 1.00 |
| CDP ever | -0.13*** | 0.05 | 0.05 | 0.05 | 0.14 | 0.99 | 0.24 | 0.18 | 0.98 |
| % low bw -1970 | 0.25** | 0.10 | 0.05 | 0.12 | 0.28 | 0.99 | -0.05 | 0.29 | 1.00 |
| Infant mortality rate-1970 | 0.15 | 0.11 | 0.18 | -0.26 | 0.19 | 0.90 | -0.10 | 0.16 | 0.98 |
| % teen motherhood -1970 | 0.21*** | 0.06 | 0.01 | 0.04 | 0.17 | 0.99 | 0.07 | 0.16 | 0.99 |
| % first born -1970 | 0.20** | 0.09 | 0.09 | -0.28* | 0.16 | 0.75 | -0.46*** | 0.14 | 0.98 |

Notes: Sample corresponds to all counties in our analyses. All county characteristics are standardized using z-scores. Mean differences correspond to the average characteristic between counties with and without WIC. The characteristics come from the City and County Data Book for 1970 and 1960. 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:** Testing for Pre-trends - Multiple Hypothesis Correction by Holm (1979)

| | Sun and Abraham Estimation - Joint Test for Pre-trends | |
|---|---|---|
| **Outcome** | **p-value from individual regressions** | **p-value after MHT correction by Holm** |
| **Panel A: Black Mothers** | | |
| Low Birth Weight | 0.104 | 0.528 |
| Very Low Birth Weight | 0.548 | 0.548 |
| Small for Gestational Age | 0.088 | 0.528 |
| Preterm | 0.140 | 0.528 |
| Birth Weight | 0.099 | 0.528 |
| High Birth Weight | 0.040 | 0.278 |
| Large for Gestational Age | 0.202 | 0.528 |
| Kessner Index for Adequate Prenatal Care | 0.004 | **0.033** |
| **Panel B: White Mothers** | | |
| Low Birth Weight | 0.001 | **0.005** |
| Very Low Birth Weight | 0.011 | 0.068 |
| Small for Gestational Age | 0.891 | 1.000 |
| Preterm | 0.256 | 0.767 |
| Birth Weight | 0.019 | 0.093 |
| High Birth Weight | 0.931 | 1.000 |
| Large for Gestational Age | 0.162 | 0.649 |
| Kessner Index for Adequate Prenatal Care | 0.000 | **0.0000** |

Notes: Column 1 of this table reports the p-values shown inside Figures 3 and 4 from a joint test of statistical significance for all pre-period coefficients. In Column 2, we report p-values for a joint significance test of all pre-period coefficients after taking into account testing for multiple outcomes using a step-down procedure introduced by Holm (1979).

**Table A.6:** de Chaisemartin and D'Haultfœuille (2020): Regression of TWFE weights ($W_{gt}$) on Factors Associated with The Treatment Effect

| | Correlation | Coef. | S.E | T-stat |
|---|---|---|---|---|
| Month_year | -0.737 | -17.094 | 0.428 | -39.962 |
| Black | -0.022 | -0.003 | 0.001 | -4.767 |
| State UI compensation per capita | 0.285 | 33.305 | 4.638 | 7.181 |
| Medicare benefits per capita | -0.282 | -37.678 | 3.424 | -11.003 |
| Hospitals per capita | 0.077 | 0.0018 | 0.0002 | 11.340 |

Notes: The outcome data are from the Vital Statistics Natality Files that cover the years between 1968 and 1981. This table reports results using a diagnostic test introduced by de Chaisemartin and D'Haultfœuille (2020).

**Table A.7:** The Effect of WIC on Main Birth Outcomes 1968–1981 by Race-County-Year of Second Trimester Cells: Sun and Abraham (2021) Estimates, Propensity Score Specification for 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 | -0.4556 | 0.0185 | -0.7245 | 0.8372 | 21.2128** | -0.1143 | 0.9993 | 0.0269 |
| | (0.5372) | (0.1558) | (0.6542) | (0.8167) | (9.9110) | (0.1262) | (0.7358) | (0.0244) |
| Control Y-mean | 7.658 | 1.122 | 11.032 | 9.283 | 3194.845 | 0.671 | 5.797 | 0.230 |
| % Effect (Coef/Control Mean) | -5.950 | 1.650 | -6.567 | 9.018 | 0.664 | -17.036 | 17.237 | 11.702 |
| Adj. R-Squared | 0.021 | 0.012 | 0.016 | 0.031 | 0.050 | 0.022 | 0.030 | 0.536 |
| Obs | 78100 | 78100 | 60780 | 60780 | 78100 | 78100 | 60780 | 65433 |
| 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 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 untreated cells are weighted by the number of births in the cell (treatment group) and $\hat{p}/(1-\hat{p})$ times the number of births in the cell (control group). The outcome data are from Vital Statistical Natality Files that cover the years between 1968 and 1981. 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. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

**Table A.8:** Robustness to Specification: The Effect of WIC on Birth Weight 1968–1981 by Race-County-Month and Year of Second Trimester Cells: Sun and Abraham (2021) and TWFE Estimates, All Counties

| | Sun and Abraham (2021) | | | | | | TWFE |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| *A: Black Mothers* | | | | | | | |
| WIC-month year-trim2 | 12.7268*** | 11.2050*** | 11.1644*** | 9.9877*** | 5.0715** | 9.2409*** | 11.7671*** |
| | (3.8047) | (2.3141) | (2.3496) | (2.2496) | (2.5363) | (2.0849) | (3.5684) |
| Control Y-mean | 3080 | 3080 | 3080 | 3080 | 3080 | 3080 | 3080 |
| Adj. R-Squared | 0.072 | 0.076 | 0.076 | 0.076 | 0.079 | 0.083 | 0.072 |
| Obs | 205159 | 205159 | 205159 | 205123 | 205123 | 205123 | 205159 |
| N States | 48 | 48 | 48 | 48 | 48 | 48 | 48 |
| N Counties | 2190 | 2190 | 2190 | 2190 | 2190 | 2190 | 2190 |
| *B: White Mothers* | | | | | | | |
| WIC-month year-trim2 | 4.8842** | 3.9324*** | 3.7819*** | 4.2434*** | 3.2032*** | 1.7173* | 4.3995** |
| | (2.0338) | (1.0630) | (1.0867) | (1.0693) | (1.0678) | (0.9092) | (1.9782) |
| Control Y-mean | 3340 | 3340 | 3340 | 3340 | 3340 | 3340 | 3340 |
| Adj. R-Squared | 0.264 | 0.285 | 0.285 | 0.285 | 0.287 | 0.290 | 0.264 |
| Obs | 474383 | 474383 | 474383 | 474335 | 474335 | 474335 | 474383 |
| N States | 48 | 48 | 48 | 48 | 48 | 48 | 48 |
| N Counties | 2886 | 2886 | 2886 | 2886 | 2886 | 2886 | 2886 |
| 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 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 untreated cells are weighted by the number of births in the cell (treatment group) and $\hat{p}/(1-\hat{p})$ times the number of births in the cell (control group). The outcome data are from Vital Statistical Natality Files that cover the years between 1968 and 1981. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level. 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.9:** Robustness to Pscore Weighting: The Effect of WIC on Main Birth Outcomes 1968–1981 by Race-County-Month and Year of Second Trimester Cells: Sun and Abraham (2021) Estimates, All and Balanced Counties

| | Low Birth Weight | | | | Birth Weight | | | |
|---|---|---|---|---|---|---|---|---|
| | All Counties-Pscore*births | All Counties-Pscore | Balance Counties-Pscore*births | Balanced Counties-Pscore | All Counties-Pscore*births | All Counties-Pscore | Balance Counties-Pscore*births | Balanced Counties-Pscore |
| *A: Black Mothers* | | | | | | | | |
| WIC-month year-trim2 | -0.3685*** | -0.1516 | -0.3773*** | -0.5225** | 12.7268*** | 7.4939 | 11.5349*** | 16.8113** |
| | (0.1249) | (0.2386) | (0.1379) | (0.2349) | (3.8047) | (5.3694) | (4.0342) | (6.9219) |
| Control Y-mean | 13 | 13 | 14 | 14 | 3080 | 3080 | 3076 | 3076 |
| % Effect (Coef/Control Mean) | -2.744 | -1.128 | -2.782 | -3.853 | 0.413 | 0.243 | 0.375 | 0.547 |
| Adj. R-Squared | 0.025 | 0.015 | 0.050 | 0.028 | 0.072 | 0.039 | 0.129 | 0.073 |
| Obs | 205159 | 205159 | 76440 | 76440 | 205159 | 205159 | 76440 | 76440 |
| N States | 48 | 48 | 36 | 36 | 48 | 48 | 36 | 36 |
| N Counties | 2190 | 2190 | 455 | 455 | 2190 | 2190 | 455 | 455 |
| *B: White Mothers* | | | | | | | | |
| WIC-month year-trim2 | -0.0481 | 0.0723 | -0.1905** | -0.5649*** | 4.8842** | 4.7834** | 4.4602* | 13.8060*** |
| | (0.0414) | (0.0916) | (0.0792) | (0.1502) | (2.0338) | (2.2830) | (2.2727) | (3.7298) |
| Control Y-mean | 6 | 6 | 7 | 7 | 3340 | 3340 | 3329 | 3329 |
| % Effect (Coef/Control Mean) | -0.749 | 1.126 | -2.910 | -8.632 | 0.146 | 0.143 | 0.134 | 0.415 |
| Adj. R-Squared | 0.051 | 0.019 | 0.099 | 0.036 | 0.264 | 0.089 | 0.432 | 0.135 |
| Obs | 474383 | 474383 | 76440 | 76440 | 474383 | 474383 | 76440 | 76440 |
| N States | 48 | 48 | 36 | 36 | 48 | 48 | 36 | 36 |
| N Counties | 2886 | 2886 | 455 | 455 | 2886 | 2886 | 455 | 455 |
| 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 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 untreated cells are weighted by the number of births in the cell (treatment group) and $\hat{p}/(1-\hat{p})$ times the number of births in the cell (control group). The outcome data are from Vital Statistical Natality Files that cover the years between 1968 and 1981. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level. 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 for 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.10:** Robustness to Timing of Treatment Assignment: The Effect of WIC on Main Birth Outcomes 1968–1981 by Race, Individual Level Data, Sun and Abraham (2021) Estimates

| | Black Mothers | | | | | White Mothers | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Pref: 30.5 days, 40 weeks 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 weeks to miss | Pref: 30.5 days, 40 weeks 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 weeks to miss |
| *A: Fraction below 2,500 Grams* | | | | | | | | | | |
| WIC-month year-trim2 | -0.3685*** | | | | | -0.0481 | | | | |
| | (0.1249) | | | | | (0.0413) | | | | |
| WIC-month year-trim2 | | -0.3695*** | | | | | -0.0509 | | | |
| | | (0.1252) | | | | | (0.0413) | | | |
| WIC-month year-trim2 | | | -0.3685*** | | | | | -0.0481 | | |
| | | | (0.1249) | | | | | (0.0413) | | |
| WIC-month year-trim2 | | | | -0.3843** | | | | | -0.0485 | |
| | | | | (0.1880) | | | | | (0.0492) | |
| WIC-month year-trim3 | | | | | -0.3606*** | | | | | -0.0421 |
| | | | | | (0.1323) | | | | | (0.0413) |
| 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 | 0.001 | 0.001 |
| Obs | 6217289 | 6225468 | 6217289 | 3978152 | 6106325 | 35845117 | 35898359 | 35845117 | 25058418 | 35158318 |
| *B: Birth Weight (in Grams)* | | | | | | | | | | |
| WIC-month year-trim2 | 12.7268*** | | | | | 4.8842** | | | | |
| | (3.8025) | | | | | (2.0331) | | | | |
| WIC-month year-trim2 | | 12.6936*** | | | | | 4.8755** | | | |
| | | (3.7268) | | | | | (2.0426) | | | |
| WIC-month year-trim2 | | | 12.7268*** | | | | | 4.8842** | | |
| | | | (3.8025) | | | | | (2.0331) | | |
| WIC-month year-trim2 | | | | 13.0333** | | | | | 4.8501** | |
| | | | | (5.3837) | | | | | (2.4147) | |
| WIC-month year-trim3 | | | | | 12.3870*** | | | | | 5.1051** |
| | | | | | (3.7772) | | | | | (2.0230) |
| Y-mean | 3080.185 | 3080.305 | 3080.185 | 3096.475 | 3079.756 | 3339.554 | 3339.728 | 3339.554 | 3353.334 | 3338.673 |
| Adj. R-Squared | 0.005 | 0.005 | 0.005 | 0.005 | 0.005 | 0.009 | 0.009 | 0.009 | 0.009 | 0.009 |
| Obs | 6217289 | 6225468 | 6217289 | 3978152 | 6106325 | 35845117 | 35898359 | 35845117 | 25058418 | 35158318 |
| 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 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 days for the average days in a month (Column 2 and 7), assigning 39 weeks for the missing gestational age (Columns 3 and 8), or keeping the 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 third trimester. The outcome data are from Vital Statistical Natality Files that cover the years between 1968 and 1981. We compile data on WIC timing from multiple sources between 1974 and 1980. Standard errors are clustered at the county level. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

**Table A.11:** Robustness Check: Using "Year" to Assign WIC Exposure

| | 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_exposure_year_trim2 | -0.3200** | | | 12.2438*** | | |
| | (0.1501) | | | (3.9467) | | |
| wic_exposure_year_trim3 | | -0.3098** | | | 11.8873*** | |
| | | (0.1548) | | | (3.9133) | |
| WIC-year | | | -0.2942* | | | 11.6744*** |
| | | | (0.1578) | | | (4.0009) |
| Control Y-mean | 13 | 13 | 14 | 3080 | 3079 | 3077 |
| Adj. R-Squared | 0.137 | 0.143 | 0.148 | 0.337 | 0.338 | 0.337 |
| Obs | 24742 | 24714 | 24493 | 24742 | 24714 | 24493 |
| N States | 48 | 48 | 48 | 48 | 48 | 48 |
| N Counties | 2173 | 2169 | 2146 | 2173 | 2169 | 2146 |
| *B: White Mothers* | | | | | | |
| wic_exposure_year_trim2 | -0.0399 | | | 5.1975*** | | |
| | (0.0416) | | | (1.9574) | | |
| wic_exposure_year_trim3 | | -0.0315 | | | 5.3736*** | |
| | | (0.0428) | | | (1.9160) | |
| WIC-year | | | -0.0341 | | | 5.1684*** |
| | | | (0.0423) | | | (1.9277) |
| Control Y-mean | 6 | 6 | 7 | 3338 | 3337 | 3335 |
| Adj. R-Squared | 0.353 | 0.345 | 0.361 | 0.764 | 0.762 | 0.764 |
| Obs | 40363 | 40399 | 40324 | 40363 | 40399 | 40324 |
| N States | 48 | 48 | 48 | 48 | 48 | 48 |
| N Counties | 2886 | 2886 | 2886 | 2886 | 2886 | 2886 |

Notes: Table reports the estimated coefficients and standard errors using 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 untreated cells are weighted by the number of births in the cell (treatment group) and $\hat{p}/(1-\hat{p})$ times the number of births in the cell (control group). The outcome data are from Vital Statistical Natality Files that cover the years between 1968 and 1981. We compile data on WIC timing from multiple sources between 1974 and 1980. The analysis is run using data at the cell level (defined 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.12:** Additional Robustness to Trimming Pscore and Excluding CSFP controls: The Effect of WIC on Main Birth Outcomes 1968–1981 by Race: Sun and Abraham (2021) Estimates

| | Low Birth Weight | | | Birth Weight | | |
|---|---|---|---|---|---|---|
| | Preferred-Pscore No Trim Include CSFP ever control | Pscore Trim Include CSFP ever control | Pscore No Trim Drop CSFP ever control | Preferred-Pscore No Trim Include CSFP ever control | Pscore Trim Include CSFP ever control | Pscore No Trim Drop CSFP ever control |
| *A: Black Mothers* | | | | | | |
| WIC-month year-trim2 | -0.3685*** | -0.3226** | -0.2673** | 12.7268*** | 11.5675*** | 10.9210*** |
| | (0.1249) | (0.1428) | (0.1121) | (3.8047) | (4.3041) | (3.5238) |
| Control Y-mean | 13 | 13 | 13 | 3080 | 3082 | 3080 |
| Adj. R-Squared | 0.025 | 0.022 | 0.034 | 0.072 | 0.065 | 0.088 |
| Obs | 205159 | 182759 | 205159 | 205159 | 182759 | 205159 |
| N States | 48 | 47 | 48 | 48 | 47 | 48 |
| N Counties | 2190 | 1987 | 2190 | 2190 | 1987 | 2190 |
| Min Cell Size | 1 | 1 | 1 | 1 | 1 | 1 |
| Max Cell Size | 3046 | 3046 | 3046 | 3046 | 3046 | 3046 |
| *B: White Mothers* | | | | | | |
| WIC-month year-trim2 | -0.0481 | -0.0352 | -0.0621 | 4.8842** | 3.7663* | 5.3025** |
| | (0.0414) | (0.0452) | (0.0527) | (2.0338) | (2.1125) | (2.2067) |
| Control Y-mean | 6 | 6 | 6 | 3340 | 3343 | 3340 |
| Adj. R-Squared | 0.051 | 0.046 | 0.058 | 0.264 | 0.242 | 0.293 |
| Obs | 474383 | 438230 | 474383 | 474383 | 438230 | 474383 |
| N States | 48 | 47 | 48 | 48 | 47 | 48 |
| N Counties | 2886 | 2670 | 2886 | 2886 | 2670 | 2886 |
| Min Cell Size | 1 | 1 | 1 | 1 | 1 | 1 |
| Max Cell Size | 9671 | 6212 | 9671 | 9671 | 6212 | 9671 |
| C,Year FE | Y | Y | Y | Y | Y | Y |
| 1970 controls-pscore weighting | Y | Y | Y | Y | Y | Y |
| Pscore-no trim | Y | | | Y | | |
| Pscore-trimmed | | Y | | | Y | |
| Control for CSFP ever | Y | Y | | Y | Y | |

Notes: Table reports the estimated coefficients and standard errors using 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. 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. The outcome data are from Vital Statistical Natality Files that cover the years between 1968 and 1981. 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. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

**Table A.13:** Robustness Check: Dropping Counties that Serve Special Population

| | Low Birth Weight | | Birth Weight | |
|---|---|---|---|---|
| | Preferred-<br>All Programs<br>Regular and Spec. Pop. | Drop counties<br>Special Pop. Only | Preferred-<br>All Programs<br>Regular and Spec. Pop. | Drop counties<br>Special Pop. Only |
| *A: Black Mothers* | | | | |
| WIC-month year-trim2 | -0.3685*** | -0.3731*** | 12.7268*** | 12.7788*** |
| | (0.1249) | (0.1251) | (3.8047) | (3.8142) |
| Control Y-mean | 13 | 13 | 3080 | 3080 |
| Adj. R-Squared | 0.025 | 0.025 | 0.072 | 0.073 |
| Obs | 205159 | 202833 | 205159 | 202833 |
| N States | 48 | 48 | 48 | 48 |
| N Counties | 2190 | 2147 | 2190 | 2147 |
| | | | | whitebirthwt2 |
| *B: White Mothers* | | | | |
| WIC-month year-trim2 | -0.0481 | -0.0498 | 4.8842** | 4.9370** |
| | (0.0414) | (0.0417) | (2.0338) | (2.0515) |
| Control Y-mean | 6 | 6 | 3340 | 3339 |
| Adj. R-Squared | 0.051 | 0.051 | 0.264 | 0.264 |
| Obs | 474383 | 462962 | 474383 | 462962 |
| N States | 48 | 48 | 48 | 48 |
| N Counties | 2886 | 2817 | 2886 | 2817 |

Notes: Table reports the estimated coefficients and standard errors using 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 untreated cells are weighted by the number of births in the cell (treatment group) and $\hat{p}/(1-\hat{p})$ times the number of births in the cell (control group). The outcome data are from Vital Statistical Natality Files that cover the years between 1968 and 1981. 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. $*$ p<0.10, $**$ p<0.05, $***$ p<0.01.

**Table A.14:** 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, Bitler et al. drops Virginia from their analysis. REIS controls+FS exposure include controls for per capita government transfers, including retirement and disability, medical care, cash public assistance, income per capita, and Food Stamp exposure. 1970 county controls include the percentage of the nonwhite, poor, urban, young, and old age population, and 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.