

NBER WORKING PAPER SERIES

THE EARNED INCOME TAX CREDIT AND INFANT HEALTH REVISITED

Daniel L. Dench
Theodore J. Joyce

Working Paper 26476
<http://www.nber.org/papers/w26476>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2019

We thank Robert Kaestner, Michael Grossman, Kitt Carpenter, David Jaeger, Bill Evans and Hilary Hoynes for comments and Tim Moore for data on homicide rates. We are grateful to seminar participants at Baruch College's Marx School of Public and International Affairs and Queen's University Belfast, the CUNY Graduate Center and the NBER Summer Institute. We have no conflicts of interest associated with this manuscript. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2019 by Daniel L. Dench and Theodore J. Joyce. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Earned Income Tax Credit and Infant Health Revisited
Daniel L. Dench and Theodore J. Joyce
NBER Working Paper No. 26476
November 2019
JEL No. H24,I38,J13

ABSTRACT

Hoynes, Miller and Simon (2015), henceforth HMS, report that the national expansion of the Earned Income Tax Credit (EITC) is associated with decreases in low birth weight. We question their findings. HMS's difference-in-differences estimates are unidentified in some comparisons, while failed placebo tests undermine others. Their effects lack a plausible mechanism as the association between the EITC and prenatal smoking also fails placebo tests. We contend that the waning of the crack epidemic is a possible confound, but we show that any number of policies directed at poor women also eliminate the effect of the EITC when aggregated to the national level. Identifying small, causal effects of a national policy at a single point in time is exceedingly challenging.

Daniel L. Dench
The Graduate Center, CUNY
365 Fifth Avenue
New York, NY 10016
ddench@gradcenter.cuny.edu

Theodore J. Joyce
Baruch College, CUNY & NBER
5 Hanover Square, 16th Floor
Suite 1602
New York, NY 10004
theodore.joyce@baruch.cuny.edu

In an award-winning article, Hoynes, Miller and Simon (2015, henceforth HMS) use variation in the Earned Income Tax Credit (EITC) by family size pre and post the 1993 expansion to test the effect of an income transfer to working families on birth outcomes. The question of whether income transfers can improve birth outcomes is of great importance to U.S. social policy. First, preterm birth (< 37 weeks gestation) and low birth weight (< 2500 grams) are the most important predictors of infant mortality, a widely used index of a nation's health. Second, there is a clear inverse gradient between adverse birth outcomes and socio-economic status. Third, the long-term effects of low birth weight on adult health appear significant (Almond and Currie 2011). With expenditures of over 66 billion dollars and 27 million recipients in 2016, the EITC is considered a highly successful anti-poverty program that uses refundable tax credits to augment the earnings of low-income workers (Tax Policy Center 2019). Given the small or non-existent federal tax liability of low earners, the EITC results in a substantial income transfer. If the EITC improves infant health, in addition to encouraging employment, then its potential welfare-enhancing impact would be even greater.

HMS follow a large literature that evaluates the EITC on primarily employment by using a difference-in-difference (DD) design to compare the birth outcomes of single, less educated women who are having a first child (parity 1) with those having second (parity 2) and higher order births (parity 3+) [Eissa and Liebman 1996; Meyer and Rosenbaum 2000; Meyer 2002; Eissa and Hoynes 2004; Eissa and Nichols 2005].¹ The largest increase in the EITC occurred with Omnibus Budget Reconciliation Act of 1993 (OBRA93). HMS contrast the change in low birth weight from 1991-1993, their pre-period, to 1994-1998 the post period. They find that the change in the generosity of the EITC after 1993 lowers

¹ In obstetrics parity is the number of previous live births. A woman giving birth for the first time would be parity zero or nulliparous. HMS define parity as the current plus previous births. Thus, HMS define a first birth as parity 1, a second birth as parity 2 and higher order births as parity 3+. We follow their definition to avoid confusion.

the incidence of low birth weight by 0.35 percentage points over a mean of 10.2 percent or a decline of 9 grams over a mean of 3206 grams by comparing single women with children to those without.²

The validity of HMS's difference-in-differences (DD) design depends on isolating a period prior to the 1993 EITC in which both the "treated" and comparison group are unexposed to the policy. A second important condition is the parallel trend assumption in which the time path of the outcome is the same for both the "treated" and comparison groups in the period prior to the intervention. Neither condition holds when HMS compare women of parity 2 or 2+ to women of parity 1. By contrast, the comparison between women of parity 2 to women of parity 3 or higher (parity 3+) provides, ostensibly, a more credible design.³ As can be seen in Figure 1, there is a large divergence in the available tax credit between women of parity 2 and parity 3+ after 1993 and barely any before. By 1996 women of parity 3+ were eligible for \$1,283 more in tax credits in 1995 dollars than women of parity 2 (Figure 1). HMS find that the EITC is associated with a decrease in low birth weight among women of parity 3+ relative to parity 2. The finding, however, is only true for black women who make up approximately 20 percent of

² The paper by HMS is the first published analysis of the effect of the 1993 expansion of the EITC on infant health. There are a number of studies that use expansions in the state EITC that report associations with birth weight. However, state EITCs are largely irrelevant between 1991-1996. By 1993 only 4 states had a refundable tax credit (Maryland, Minnesota, Vermont and Wisconsin) while Rhode Island had a nonrefundable credit (Hotz and Scholz 2003). Strully, Rehkopf and Xuan (2010), Markowitz et al. (2017), Komro et al. (2019) use variation in state EITCs to analyze their association with birth weight. A number of results raise questions about the robustness of the findings. First, each study estimates increases in birthweight associated with the state EITC that are twice as large as HMS despite state tax credits of roughly one-tenth to one-fifth the magnitude of the federal tax credits. Second, despite the use of two-way fixed effects models, policy endogeneity is a concern as the states with EITCs tend to have more generous social welfare policies. Yet, none of the studies control for state-specific trends or use an event-study design to test for pre-trends. Lastly, each study finds effects of state EITCs on the outcomes of first births, which seems implausible. The available tax credit is between \$30-\$60 per year, and thus any effect is more indicative of a failed placebo test than a credible treatment. See also studies by Bruckner, Rehkopf and Catalano 2013; Hamad and Rehkopf (2015). For a more detailed review of studies analyzing the state EITCs, see Appendix B.

³ In fact, the level and time-series pattern of the rate of low birth weight for women of parity 5 or higher are very different from those of women of parity 2 and call into question their use in the analysis. The time-series pattern of low birth weight among women of parity 3 and 4 is much closer to that of parity 2 (see Appendix A, Figure A-1A).

EITC filers. There is no association among whites or Hispanics. As we show below, however, results for black women fail a basic placebo test (Hotz, Mullin and Scholz 2006).

HMS suggest that the effect of the EITC on low birth weight among black women can be explained, in part, by a decrease in prenatal smoking and an increase in prenatal care. Prenatal smoking is the most important modifiable risk factor for low birth weight. HMS's results, however, are too large and inconsistent to be credible. They report that prenatal smoking fell by 2.41 percentage points more among black women of parity 3+ relative to parity 2. This implies an income elasticity of -2.75 and a quit rate of over 100 percent among women induced to work by the EITC, both implausibly large. In addition, HMS's smoking results fail a placebo test. Prenatal care, on the other hand, lacks a convincing relationship to birth outcomes, but even if prenatal care were effective, HMS estimated increases in prenatal care are too small to be relevant.

So, what did cause the fall in the rate of low birth weight among black women of higher parity? We speculate that the waning of the crack-cocaine epidemic in the early 1990s may explain the particular pattern of HMS's results. As some evidence, we use New York City birth certificate data that shows large differential trends in low birth weight and prenatal drug use by race and parity from 1984-1998. Although the fading crack epidemic is a possible explanation, the larger point is that identifying small effects of a national policy change on birth outcomes at a single point in time is exceedingly challenging. To illustrate, we aggregate from the state to the national level the three state variables used by HMS to control for policies that may affect infant health. Their inclusion at the state level has little explanatory power in HMS's specification, but they greatly diminish or completely eliminate the effect of the EITC when entered at the same level of aggregation as the EITC.

I. The 1993 EITC Expansion: Parity 1 vs. Parity 2 and 2+

Federal legislation resulted in the expansion of the EITC in 1986, 1990 and 1993. The expansions in 1990 and 1993 were phased in over the subsequent three years: 1991-1993 and 1994-1996. Up until 1994, the available tax credit through the EITC was almost identical for families with one or more children (parity 2+ as specified by HMS). As shown in Figure 1, the available tax credit for women of parity 3+ increased greatly relative to women of parity 2 beginning in 1994. In that same year, women of parity 1 became eligible for a tax credit for the first time. These staggered expansions make comparisons between women of parity 2 versus parity 1 almost impossible to identify when the study period is limited to effective tax years 1991-1998 as used by HMS. For instance, from 1990 to 1993 the EITC credit increased by \$401 (\$222 + \$179 in 1995 \$) for women of parity 2 relative to parity 1 which *exceeded* the post-period increase of \$264 (\$578 - \$314 in 1995 \$) from 1994-1998.⁴ In other words, the increase in the EITC for HMS's treatment group, women of parity 2, relative to their counterfactual, women of parity 1, is larger from 1990 to 1993 than in their HMS's post period. To lessen the impact of the 1990 expansion of the EITC on their treatment group, HMS limit the pre-period to 1991-1993. The increase over this two-year pre-period at \$179 is still 68 percent of the post-EITC increase. Yet limiting the pre-period to 1991-1993 does little to mitigate the cumulative exposure of their treatment group to increases in the EITC during their pre-period. The reason is because HMS's mapping from the child's birth date to the year the mother is exposed to the EITC, what HMS refer to as the woman's "effective tax year," is an approximation at best.⁵ For instance, HMS assume that women who gave birth in

⁴ As shown in Figure 1, the EITC increases by \$578 between 1994 and 1998 for women of parity 2 but the EITC increases by \$314 for women of parity 1 over the same period. The net increase is \$264 (\$578-\$314).

⁵ One reason for HMS's specific mapping from the effective tax year to the month of birth (see their HMS Table 1) is an effort to time the tax credit to the third trimester of pregnancy. Given the crudeness of the data, such fine-tuning of the EITC exposure is unlikely to be accurate and is one more reason to use 1987-1990 as the pre-period during which there were no changes in the EITC. Second, focus on the third trimester is too restrictive as it is not

January through August of 1992 were only eligible for the 1990 EITC. But many of these women were likely working through much, if not all, of 1991 and therefore received the higher 1991 tax credit in February or March of 1992. We agree with HMS that mapping a pregnancy to a specific EITC year is not precise, but it's also unnecessary. The years 1987-1990 provide a superior pre-period with 1991-1998 as the more appropriate post period (see Figure 1). Specifically, there is no nominal change in the EITC from 1987 to 1990 between women of parity 2 or 2+ versus parity 1. There is a continuous relative increase from 1991 to 1994 of \$665 in 1995 dollars for women of parity 2 versus parity 1 that remains nominally unchanged from 1995-1998. Indeed, the \$665 increase in available tax credits for women of parity 2 compared to parity 1 is the largest relative expansion in a 4-year period between the two groups over the entire existence of the EITC.

To illustrate, we first replicate HMS's DD regression estimates comparing the rate of low birth weight among women of parity 2 relative to parity 1 using 1991-93 as the pre-period and 1994-1998 as the post period. HMS aggregate individual-level birth certificate data for low birth weight into 47,687 non-zero cells for single women with no more than a high school education defined by state, year, parity, maternal education, race, ethnicity and age from 1981 to 1998 and estimate the following regression (HMS, Appendix B).

$$(1) Y_{pjst} = \alpha + \delta(After_t * Treat_p) + \pi X_{st} + \rho_p + \varphi_j \varphi_k + \delta_t + \tau_s + \epsilon_{pjst}$$

Y_{pjst} is the rate of low birth by parity p , demographic group j with interactions of demographic group k , state s and year t . \mathbf{X} is the set of state policies: Medicaid/SCHIP, welfare reform and the state

the only point in a pregnancy that could benefit from an intervention. Neural tubal defects, which occur early in pregnancy, are vulnerable to smoking and drinking. Precursors of preterm birth such as vaginal bleeding, preeclampsia and gestational diabetes should be diagnosed and treated whenever they arise during pregnancy. This explains the current emphasis on pre-conception health (see <https://www.acog.org/Clinical-Guidance-and-Publications/Committee-Opinions/Committee-on-Gynecologic-Practice/The-Importance-of-Preconception-Care-in-the-Continuum-of-Womens-Health-Care>).

unemployment rate that vary by year. $Treat_p$ equals one for the parities experiencing expansion in available tax credits (i.e., parity 2) relative to the controls (i.e., parity 1). HMS limit the sample to single mothers with high school or less of completed schooling. We report the DD coefficient, δ , in the top panel of Table 1. Among all women, an increase in the EITC beginning in 1994 is associated with a 0.164 percentage point decline in the rate of low birth weight among women of parity 2 relative to parity 1 (HMS, Table 2). In Panel B, we use the same pre-period as HMS but we drop California, New York, Texas and Washington because they did not report mother's education prior to 1989. The differences in the estimates between Panel A and B are modestly more supportive of HMS. In the bottom panel of Table 1, we show results from running the same specification but using 1987-1990 as the pre-period and 1991-1998 as the post-period. The protective effect of the EITC using this pre-period is more than twice as large for all women (-0.564), black women (-0.920) and white women (-0.352) than using 1991-1993 as the pre-period for the same groups (Table 1, Panel c). We estimate the same specification comparing women of parity 2+ to women of parity 1. The results are qualitatively the same as parity 2 versus parity 1 (Table 1, Panel D).

Estimating the effect of the EITC with a more coherent pre- and post-period appears to further support HMS's conclusions. A key assumption of the DD, however, is that trends between the "treated" and comparison group leading up to intervention should be parallel. As evidence, HMS estimate an event-study specification that shows leads and lags of the estimated DD.⁶ We replicate their specification (Figure 2, Panel A) but then extend the data back to 1987 (Figure 2, Panel B). The key point is that the rate of low birth weight among women of parity 2 relative to parity 1 declines continuously from 1987 to 1998 (Figure 2, Panel B) with no obvious discontinuity after 1990 or 1993. Large

⁶ (2) $Y_{pjst} = \alpha + \sum_{k=-2}^5 [\delta_k (\mathbf{1}(\mathbf{After}_{(k=t-t^*)} * Treat_p))] + \pi X_{st} + \rho_p + \varphi_j \varphi_k + \delta_t + \tau_s + \epsilon_{pjst}$

where $\mathbf{1}$ is the indicator function and $t^* = 1993$ and $t = 1991, \dots, 1998$

differential trends in low birth weight are also evident for women of parity 2+ relative to women of parity 1 (Figure 2, Panels C and D). An F-test decisively rejects the null that coefficients on the interactions from 1987-1992 are different from zero for both parity 2 relative to parity 1 ($F_{6,46}=7.85$), and 2+ relative to 1 ($F_{6,46}=5.52$).

We contend that it is not possible to isolate the effect of the 1993 EITC on the birth outcomes of women of parity 2 or parity 2+ relative to parity 1 between 1991 and 1998. Two key conditions for a valid DD are violated. The continuous increase in available tax credits from 1991 to 1994 for women of parity 2 and parity 2+ relative to parity 1 means the “treatment group” was exposed to the intervention in the pre-period. This could have been rectified by using 1987-1990 as the pre-period and 1991-1998 as the post period. However, women of parity 1 are a poor counterfactual for women of higher parity over this period. Differential trends in low birth weight between women of parity 2+ relative to women of parity 1 from 1987-1992 undermine this design. We turn next to the comparisons between women of parity 3+ versus parity 2 that meet these two key conditions: similar exposure to the EITC tax credit between the “treated” and comparison groups in the pre-period and no differential trends in low birth weight leading up to the 1993 expansion.

II. The 1993 EITC Expansion: Parity 2 vs. Parity 3+

Numerous studies of the EITC have used the large increases in the available tax credits to women of parity 3+ relative to women of parity 2 after 1993 to identify the effects of the EITC on employment (Figure 1).⁷ Between 1993-1996, the EITC rose by \$1,861 for women of parity 3+ and by \$578 for women of parity 2, a net difference of \$1,283 in 1995 dollars. HMS use equation (1) to

⁷ See Meyer 2002; Eissa and Hoynes 2004; Eissa and Nichols 2005; Hotz, Mullin and Scholz 2006; Evans and Garthwaite 2014).

estimate the difference in birth outcomes between women of parity 3+ versus parity 2 between 1991-1998. The pre-period in this design, 1991-93, is not confounded by differential exposure to the EITC of any substance, unlike the comparison for women between parity 2 and parity 1.

The results provide support for HMS's conclusion that the 1993 EITC expansion improved birth outcomes for low-income, single women of parity 3+ relative to women of parity 2. In Table 2, panel A, we replicate HMS's results for all women and then by race/ethnicity (HMS Tables 2 and 3). Women of parity 3+ experience declines of 0.347 percentage points more than women of parity 2, a 3 percent decline given a mean rate of low birth weight of 10.2 percent. The results for all women, however, are driven by birth outcomes among black women for whom rates of low birth weight fell 0.715 percentage points more among women of parity 3+ relative to parity 2. Importantly, there is no association between the EITC and rates of low birth weight among white or Hispanic women. The point estimate for white women is so close to zero that even the upper 95 percent confidence interval suggests a clinically inconsequential effect.

The large increase in the 1993 EITC for women of parity 3+ provides a natural placebo test (Hotz, Mullin and Scholz 2006). There should be no differential effect of the EITC on women of parity 3 versus those of parity 4 or parity 4+ as they all experience the same increase in the available tax credits. HMS report this test for all women in their Appendix, which we have replicated in column (1) of Table 2, Panels B, C, and D. Specifically, the rate of low birth weight declines by 0.225 percentage points more among women of parity 3 versus those of parity 2 after the 1993 expansion in the EITC. The effect is slightly larger between women of parity 4 versus parity 3 (-0.240) and even larger for women of parity 4+ versus those of parity 3 (-0.261).⁸ HMS acknowledge that failure of this placebo test suggests

⁸ The large differential effects between women of parity 4 and 4+ versus three are all the more surprising because the higher parity women have to distribute the same tax credit over a larger family, which could diminish any protective effect of income on the most recent birth.

differential trends in low birth weight by parity. What HMS don't point out is that only the estimates for black women fail this test (Table 2, column 2, Panels B, C, and D). There is no differential effect of the EITC on white or Hispanic women of parity 3 versus parity 4 or parity 4+, which is consistent with the null findings for both groups between women of parity 3 versus parity 2 (Table 2, columns 3 and 4).⁹

III. The Etiology of Low Birth Weight and Plausible Mechanisms

Low birth weight (infants less than 2,500 grams) consists of infants born preterm (less than 37 weeks gestation) and/or those who were growth restricted. In 2015, 69 percent of all low birth weight births were preterm. In an exhaustive review of the literature on preterm birth, the Institute of Medicine (IOM) states in the abstract, "The current methods for diagnosis and treatment of preterm labor are currently based on an inadequate literature, and little is known how preterm birth can be prevented" (Institute of Medicine 2007, p.2).¹⁰ Growth retardation, on the other hand, has a number of antecedents that include smoking, small stature, low weight gain, and first births with smoking being the most important (Kramer et al. 2000). HMS propose two mechanisms by which the EITC might protect against adverse birth outcomes: prenatal care and prenatal smoking.¹¹ We find prenatal care to be an implausible explanation. As we detail in Appendix A, Section III, the association between prenatal care and birth outcomes from observational studies is modest at best and vulnerable to selection bias and differential measurement error. Reviews of randomized trials of prenatal or augmented prenatal care find little or no effect on birth outcomes (Corman, Dave and Reichman, forthcoming). Moreover, HMS's

⁹ The failed placebo tests among black women do not reflect differential trends in low birth weight among women of higher parity (see the Appendix A, Section II).

¹⁰ See also Iams (2014) for a more recent review.

¹¹ Our discussion focuses on the sample of black women as there is little consistent evidence of an association between the EITC and birth outcomes for either white or Hispanic women.

estimates of the effect of the EITC on prenatal care are so small as to be inconsequential; they also fail our placebo tests.

Prenatal smoking, on the other hand, is considered the most important modifiable determinant of low birth weight (Floyd et al. 1993; U.S. Department of Health and Human Services, 2001). Based on HMS's data and specifications, the EITC lowered the prevalence of smoking by 2.41 percentage points among black women of parity 3+ relative to parity 2 over a mean prevalence of 19 percent. The finding is implausible for several reasons. First, the implied quit rate among newly employed women necessary to generate a 2.41 percentage point drop in the prenatal smoking of all pregnant women is over 100 percent.¹² Second, workplace smoking bans could not achieve such large reductions. The best estimate is that workplace smoking bans reduce smoking by 5 percentage points (Evans, Farrelly and Montgomery 1999). Based on the increase in employment in HMS's sample, this would reduce smoking by 0.35 percentage points.¹³ Third, HMS's estimates assume smoking is an inferior good among low income, single women with an implied income elasticity of -2.75, a huge response that is also outside the bounds of the literature. A meta-analysis by economists suggests that, on balance, smoking is a normal good (Gallet and List 2003). Pregnant women are more likely to quit than non-pregnant women, but the reasons women give for quitting are the fear of adverse birth outcomes, medical advice and

¹² Consider 100,000 working women. Assume 20 percent or 20,000 women are smokers. Assume 10 percent or 10,000 are new workers induced by the EITC based on HMS estimates of the increase in labor force participation among black women (HMS Appendix Table 1). Assume 20 percent or 2,000 of the 10,000 new workers are smokers. Assume all new workers quit smoking. Thus, the total number of smokers in the population falls from 20,000 to 18,000 and the smoking prevalence falls from 20 percent to 18 percent or 2 percentage points. To put a quit rate of 100 percent into perspective, a Cochrane review of 72 randomized trials of prenatal smoking cessation programs reports quit rates of 6 percent among all trials that were reviewed, but only 3 percent among those with the strongest designs (Lumley et al. 2014)

¹³ Again consider a population of 100,000 black women of whom 20 percent or 20,000 smoke. Ten thousand enter the workforce of whom 2000 smoke. Based on Evans, Farrelly, Montgomery (1999), 7000 women would end up working at sites with workplace bans of which 1400 smoke (0.20×7000). Bans reduce smoking by 5 percentage points or by 350 ($.05 \times 7000$). As a result, smoking in the population would fall from 20,000 to 19,650 or from 20 percent to 19.65 percent—a decline of 0.35 percentage points.

nausea (Floyd et al. 1993). HMS suggests that cash transfers from the EITC received mostly in February and March, which are spent largely on durables and transportation, induce unrealistic quit rates among women of parity 3+ relative to parity 2.

Perhaps the most important reason to doubt HMS's smoking estimates for black women is that they fail a basic placebo test. Recall that women of parity 3, 4 and 4+ all received the same increase in the EITC after 1993. There should be, therefore, no differential decline in smoking among women of parity 4 or 4+ relative to parity 3. The results in Table 3 indicate otherwise. Prenatal smoking falls 1.23 percentage points more among black women of parity 4 relative to parity 3 and 1.18 percentage points more among black women of parity 4+ relative to parity 3. As with low birth weight, the failed placebo tests point to differential trends in prenatal smoking by parity.¹⁴

IV. Alternative Explanations

Despite the lack of interventions to prevent preterm birth or credible changes in smoking associated with the EITC, trends in low birth weight reported by HMS and documented here vary significantly by race and parity.¹⁵ One plausible factor consistent with differential trends in low birth weight by race and parity is the impact of the crack-cocaine epidemic of the 1980s and 1990s. There is broad consensus among social scientists that growth in crack-cocaine markets expanded rapidly in poor urban areas between 1983 and 1990. The spread of crack led to a dramatic upsurge in homicide that

¹⁴ Despite the failed placebo tests and the lack of a convincing mechanism, HMS argue that the "preponderance of evidence" still favors a causal interpretation of their results (see footnote 8). Part of that evidence is their expanded analysis of the three EITC expansions in 1986, 1990 and 1993. In Appendix A, IV, we show that there is no association between the EITC expansions in 1986 and 1990 and even the effect of the 1993 expansion is eliminated when we allow quadratic trends of low birth weight by parity.

¹⁵ As HMS show, the EITC was not associated with changes in fertility suggesting that other determinants of growth retardation such as small stature and first births were also not changing differentially by race and parity.

was concentrated among young black males between the ages of 15-24 (Blumstein 1995; Cook and Laub 1998; Cork 1999; Blumstein, Rivara and Rosenfeld 2000; Grogger and Willis 2000; Fryer et al. 2013; Evans, Garthwaite and Moore 2016, 2018). For reasons that are not well understood, the crack epidemic affected black women more than whites and Hispanics. In the largest population prevalence study ever undertaken, the urines of 29,494 women were tested for perinatal substances at 202 California hospitals in 1992. The percent of women exposed to cocaine at delivery was 13 times greater among black, non-Hispanics (7.79 percent) than white, non-Hispanics (0.60 percent) and Hispanics (0.55 percent) [Vega et al. 1993]. Prenatal exposure to crack was also concentrated among older, higher parity black women as demonstrated in numerous studies from urban hospitals and clinics across around the country.¹⁶ In a meta-analysis of 31 studies of cocaine and birth outcomes, rates of low birth weight were 3 to 4 times greater and average birth weights were 492 grams less among women exposed relative to unexposed to cocaine prenatally (Gouin et al. 2011). Exposure to crack was highly correlated with exposure to tobacco and other risk factors for adverse birth outcomes such as alcohol and sexually transmitted infections (STIs). We are not arguing that the pharmacological impact of cocaine alone was responsible for the rise in low birth weight, but that the crack epidemic was associated with a social and behavioral breakdown in certain poor communities with devastating consequences to its residents.

There are no national data on prenatal exposure to crack-cocaine let alone by race and parity. However, the New York City Department of Health collected information on drug use during pregnancy as part of its vital registration system for births. In Appendix A, Section V we provide a descriptive presentation of trends in prenatal exposure to cocaine and low birth weight by race and parity that is

¹⁶ Hospital-based studies of prenatal cocaine exposure in numerous urban centers found the average age of users was between 25 and 29 and the average parity was 3 as computed by HMS (Zuckerman et al. 1989; Hadeed and Siegel 1989; McCalla et al. 1991; Phibbs, Bateman and Schwartz 1991; Handler et al. 1991; Bateman et al. 1993; Eyler et al. 1994; Chazotte et al. 1995; Singer et al. 2002).

consistent with our conjecture that the association between the EITC and the decline in low birth weight among black women may be related to the crack epidemic.

Although the waning of the crack epidemic is a possible explanation for why only black births of parity 3+ are associated with the EITC, the lack of data on prenatal exposure to cocaine nationally limits a convincing test. The more general point is the difficulty of identifying the effect of a national policy at a single point in time in what Manzi refers to a “dense causal environment” (Manzi 2012). For instance, HMS’s sample includes over 47,000 cells and yet, identification of the EITC derives from changes in low birth weight over nine years and two parity groups. To demonstrate, we re-estimate HMS’s equation (1) and contrast the estimated effect of the EITC on low birth weight with those from a model in which we drop the state policy variables, the state fixed effects, and all demographic fixed effects. This limited model includes only year and parity fixed effects and an interaction term. Estimates in the first column of Table 4, Panel A, replicate the results from HMS’s full model while those in column 2 adjust for only year and parity fixed effects. The coefficients are almost identical but the model in column (1) has 86 parameters and the one in column (2) has 12. The same pattern persists for black women in Panel B of Table 4.

With such limited identifying variation, any number of factors during this period could confound the effect of the EITC.¹⁷ To illustrate, we aggregate HMS’s three state-year variables to the annual level and interact them with an indicator of women of parity 3+. The three policies are the proportion of women who are exposed to welfare reform; the unemployment rate, and the income eligibility threshold for pregnant women applying for Medicaid or SCHIP as a percentage of the federal poverty

¹⁷ Kleven (2019) shows convincingly that welfare reform and the macroeconomy—not the 1993 EITC expansion—were more likely the cause of the large increase in employment among less educated single women with children from 1993 to 2000.

level. Each policy has a plausible link to low birth weight.¹⁸ In column (3) we include welfare reform interacted with parity along with the indicator of the EITC expansion; in column (4) we drop the EITC interaction. Inclusion of welfare reform lowers the coefficient on the EITC by half for all women but it completely eliminates the effect of the EITC among black women (Table 4, Panel B). In column (4) we show that welfare reform is associated with a 0.49 percentage point decline in low birth weight for all women of parity 3+ relative to parity 2 and a 1.16 percentage point decline among black women. Both estimates exceed the effect of the EITC. Inclusion of the annual unemployment rate interacted with women of parity 3+ eliminates any association between the EITC and low birth weight and suggests that a one percentage increase in the annual unemployment rate increases the rate of low birth weight between 0.14 and 0.38 for women of parity 3+ relative to parity 2 depending on race (Table 4, columns 5 and 6).¹⁹ Importantly, none of the three policy variables have any association with rate of low birth weight among black women in HMS's full model when allowed to vary by state and year. And yet, when aggregated to the national level, they appear to have a robust association with low birth weight that varies by parity. We are not claiming that welfare reform, unemployment, the Medicaid eligibility expansions or the crack epidemic have a causal association with low birth weight. Nevertheless, when aggregated to the national level, the differential associations of these factors with low birth weight by race and parity highlight the difficulty of deriving causal estimates from a national policy at a single point in time.

¹⁸ Numerous states received waivers to AFDC that permitted time limits, work requirements, and other features that formed the basis of PRWORA in 1996. See <https://aspe.hhs.gov/report/state-implementation-major-changes-welfare-policies-1992-1998> (last accessed August 4, 2019)

¹⁹ We obtain similar result when we include the homicide rate of black males 15-24 years of age at the national level as a proxy for the crack epidemic (Appendix A Section V, Table A3). Its inclusion eliminates the effect of the EITC completely.

Conclusion

We end by praising HMS for trying to test the effect of a plausibly exogenous increase in income on infant health. The EITC represents a sizable transfer to working single women with children that may improve birth outcomes (Brownell et al. 2016). In systematically reviewing their evidence, however, we conclude that claims of a causal association between the EITC and low birth weight cannot be drawn from HMS's design and data. Specifically, we show that comparisons between women of parity 2 versus parity 1 are flawed because women of parity 2 were exposed to the EITC in HMS's pre-period. Similarly, differential trends in low birth weight between women of parity 2+ relative to parity 1 confound their association. We agree that the most credible contrast is between women of parity 3+ relative to parity 2. However, we find no effect for whites or Hispanics, a conclusion supported by placebo tests. By contrast, the estimated protective effect of the EITC on low birth weight among black women of parity 3+ relative to parity 2 is undermined by failed placebo tests. Moreover, this association also lacks a credible mechanism. Not only are the estimated effects of the EITC on smoking among black women of parity 3+ relative to parity 2 too large to be plausible, but they also fail a placebo test. Finally using HMS's own policy variables, we show that they have no effect on low birth weight at the state level, but have robust associations with the rate of low birth weight by parity when aggregated to the level of the EITC. Indeed, inclusion of welfare reform and the national unemployment rate eliminate the effect of the EITC. In the end, we cannot conclude that income transfers such as the EITC have no protective effect against low birth weight. Our point is that efforts to identify small causal effects at the national level without a sharp discontinuity and plausible mechanisms are vulnerable to confounding from ongoing events and policies.

REFERENCES

- Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *The Journal of Economic Perspectives* 25 (3): 153–72.
- Bateman, D A, S K Ng, C A Hansen, and M C Heagarty. 1993. "The Effects of Intrauterine Cocaine Exposure in Newborns." *American Journal of Public Health* 83 (2): 190–93. <https://doi.org/10.2105/AJPH.83.2.190>.
- Blumstein, Alfred. 1995. "Youth Violence, Guns, and the Illicit-Drug Industry." *The Journal of Criminal Law and Criminology* (1973-) 86 (1): 10–36. <https://doi.org/10.2307/1143998>.
- Blumstein, Alfred, Frederick P. Rivara, and Richard Rosenfeld. 2000. "The Rise and Decline of Homicide—and Why." *Annual Review of Public Health* 21 (1): 505–41. <https://doi.org/10.1146/annurev.publhealth.21.1.505>.
- Brownell, M. D., M. J. Chartier, N. C. Nickel, D. Chateau, P. J. Martens, J. Sarkar, E. Burland, et al. 2016. "Unconditional Prenatal Income Supplement and Birth Outcomes." *Pediatrics* 137, no. 6 (June): e20152992–e20152992. <https://doi.org/10.1542/peds.2015-2992>.
- Bruckner, Tim A., David H. Rehkopf, and Ralph A. Catalano. "Income Gains and Very Low-Weight Birth among Low-Income Black Mothers in California." *Biodemography & Social Biology* 59, no. 2 (July 3, 2013): 141–56. <https://doi.org/10.1080/19485565.2013.833802>.
- Chazotte, Cynthia, Joan Youchah, and Margaret Comerford Freda. "Cocaine Use during Pregnancy and Low Birth Weight: The Impact of Prenatal Care and Drug Treatment." *Seminars in Perinatology* 19, no. 4 (August 1995): 293–300. [https://doi.org/10.1016/S0146-0005\(05\)80044-8](https://doi.org/10.1016/S0146-0005(05)80044-8).
- Cook, Philip J., and John H. Laub. 1998. "The Unprecedented Epidemic in Youth Violence." *Crime and Justice* 24: 27–64.
- Cork, Daniel. 1999. "Examining Space—Time Interaction in City-Level Homicide Data: Crack Markets and the Diffusion of Guns Among Youth." *Journal of Quantitative Criminology* 15 (4): 379–406.
- Corman, Hope, Dhaval Dave and Nancy Reichman. Forthcoming. "Effect of Prenatal Care on Birth Outcomes: Reconciling a Messy Literature" *Southern Economic Journal*.
- Eissa, N., and J. B. Liebman. 1996. "Labor Supply Response to the Earned Income Tax Credit." *The Quarterly Journal of Economics* 111 (2): 605–37. <https://doi.org/10.2307/2946689>.
- Eissa, Nada, and Hilary Williamson Hoynes. 2004. "Taxes and the Labor Market Participation of Married Couples: The Earned Income Tax Credit." *Journal of Public Economics* 88 (9–10): 1931–58. <https://doi.org/10.1016/j.jpubeco.2003.09.005>.
- Eissa, Nada, and Austin Nichols. 2005. "Tax-Transfer Policy and Labor-Market Outcomes." *American Economic Review* 95 (2): 88–93. <https://doi.org/10.1257/000282805774670527>.
- Evans, William N., Matthew C. Farrelly, and Edward Montgomery. 1999. "Do Workplace Smoking Bans Reduce Smoking?" *American Economic Review* 89 (4): 728–47. <https://doi.org/10.1257/aer.89.4.728>.
- Evans, William N., and Craig L. Garthwaite. 2014. "Giving Mom a Break: The Impact of Higher EITC Payments on Maternal Health." *American Economic Journal: Economic Policy* 6 (2): 258–90.
- Evans, William N., Craig Garthwaite, and Timothy J. Moore. 2016. "The White/Black Educational Gap, Stalled Progress, and the Long-Term Consequences of the Emergence of Crack Cocaine Markets." *Review of Economics and Statistics* 98 (5): 832–47.

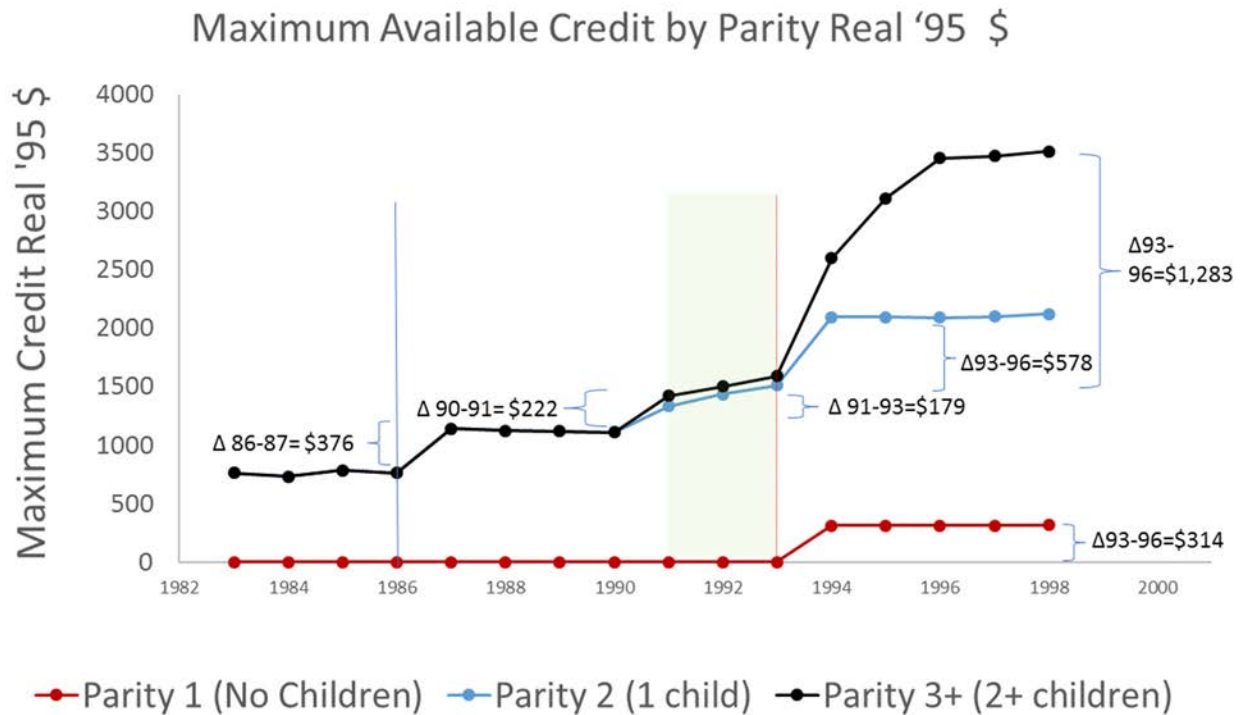
- https://doi.org/10.1162/REST_a_00635.
- Evans, William, Craig Garthwaite, and Timothy Moore. 2018. "Guns and Violence: The Enduring Impact of Crack Cocaine Markets on Young Black Males." w24819. Cambridge, MA: National Bureau of Economic Research. <https://doi.org/10.3386/w24819>.
- Eyler, Fonda Davis, Marylou Behnke, Michael Conlon, Nanci Stewart Woods, and Barbara Frentzen. 1994. "Prenatal Cocaine Use: A Comparison of Neonates Matched on Maternal Risk Factors." *Neurotoxicology and Teratology* 16 (1): 81–87. [https://doi.org/10.1016/0892-0362\(94\)90012-4](https://doi.org/10.1016/0892-0362(94)90012-4).
- Floyd, R. Louise, Barbara K. Rimer, Gary A. Giovino, Patricia D. Mullen, and Susan E. Sullivan. 1993. "A Review of Smoking in Pregnancy: Effects on Pregnancy Outcomes and Cessation Efforts." *Annual Review of Public Health* 14 (1): 379–411. <https://doi.org/10.1146/annurev.pu.14.050193.002115>.
- Fryer, Roland G., Paul S. Heaton, Steven D. Levitt, and Kevin M. Murphy. 2013. "Measuring Crack Cocaine and Its Impact." *Economic Inquiry* 51 (3): 1651–81. <https://doi.org/10.1111/j.1465-7295.2012.00506.x>.
- Gallet, Craig A., and John A. List. 2003. "Cigarette Demand: A Meta-Analysis of Elasticities." *Health Economics* 12 (10): 821–35. <https://doi.org/10.1002/hec.765>.
- Gouin, Katy, Kellie Murphy, and Prakesh S. Shah. 2011. "Effects of Cocaine Use during Pregnancy on Low Birthweight and Preterm Birth: Systematic Review and Metaanalyses." *American Journal of Obstetrics and Gynecology* 204 (4): 340.e1-340.e12. <https://doi.org/10.1016/j.ajog.2010.11.013>.
- Grogger, Jeff, and Michael Willis. "The Emergence of Crack Cocaine and the Rise in Urban Crime Rates." *The Review of Economics and Statistics* 82, no. 4 (2000): 519–29.
- Hadeed, A. J., and S. R. Siegel. 1989. "Maternal Cocaine Use during Pregnancy: Effect on the Newborn Infant." *Pediatrics* 84 (2): 205–10.
- Hamad, Rita, and David H. Rehkopf. "Poverty, Pregnancy, and Birth Outcomes: A Study of the Earned Income Tax Credit." *Paediatric and Perinatal Epidemiology* 29, no. 5 (September 2015): 444–52. <https://doi.org/10.1111/ppe.12211>.
- Handler, A., N. Kistin, F. Davis, and C. Ferré. 1991. "Cocaine Use during Pregnancy: Perinatal Outcomes." *American Journal of Epidemiology* 133 (8): 818–25.
- Hotz, V. Joseph, and John Karl Scholz. 2003. "The Earned Income Tax Credit." In Robert Moffit (ed), *Mean Tested Transfer Programs in the United States*. (Chicago: University of Chicago Press.
- Hotz, V. Joseph, Charles H. Mullin, and John Karl Scholz. 2006. "Examining the Effect of the Earned Income Tax Credit on the Labor Market Participation of Families on Welfare." Working Paper 11968. National Bureau of Economic Research. <https://doi.org/10.3386/w11968>.
- Hoynes, Hilary, Doug Miller, and David Simon. 2015. "Income, the Earned Income Tax Credit, and Infant Health." *American Economic Journal: Economic Policy* 7 (1): 172–211. <https://doi.org/10.1257/pol.20120179>.
- Iams, Jay D. "Prevention of Preterm Parturition." Edited by Caren G. Solomon. *New England Journal of Medicine* 370, no. 3 (January 16, 2014): 254–61. <https://doi.org/10.1056/NEJMcpl103640>.

- Institute of Medicine (U.S.), Richard E. Behrman, and Adrienne Stith Butler, eds. 2007. *Preterm Birth: Causes, Consequences, and Prevention*. Washington, D.C: National Academies Press.
- Kleven, Henrik. 2019. "The EITC and the Extensive Margin: A Reappraisal." Manuscript Princeton University 2019).
https://www.henrikkleven.com/uploads/3/7/3/1/37310663/kleven_eitc_sep2019.pdf (last accessed October 4, 2019)
- Komro, Kelli A., Sara Markowitz, Melvin D. Livingston, Otto Lenhart, and Alexander C. Wagenaar. 2019. "Effects of State-Level Earned Income Tax Credit Laws on Birth Outcomes by Race and Ethnicity." *Health Equity* 3(1). Accessed July 23, 2019.
<https://www.liebertpub.com/doi/10.1089/heq.2018.0061>.
- Kramer, M.S., L. Seguin, J. Lydon, and L. Goulet. "Socio-Economic Disparities in Pregnancy Outcome: Why Do the Poor Fare so Poorly?" *Paediatric and Perinatal Epidemiology* 14, no. 3 (July 2000): 194–210. <https://doi.org/10.1046/j.1365-3016.2000.00266.x>.
- Lumley, Judith, Catherine Chamberlain, Therese Dowswell, Sandy Oliver, Laura Oakley, and Lyndsey Watson. "Interventions for Promoting Smoking Cessation during Pregnancy." In *Cochrane Database of Systematic Reviews*, edited by The Cochrane Collaboration. Chichester, UK: John Wiley & Sons, Ltd, 2009.
<https://doi.org/10.1002/14651858.CD001055.pub3>.
- Manzi, Jim. 2012. Uncontrolled: The Surprising Payoff of Trial-and-Error for Business, Politics and Society. (New York: Basis Books).
- Markowitz, Sara, Kelli A. Komro, Melvin D. Livingston, Otto Lenhart, and Alexander C. Wagenaar. 2017. "Effects of State-Level Earned Income Tax Credit Laws in the U.S. on Maternal Health Behaviors and Infant Health Outcomes." *Social Science & Medicine* 194(December): 67–75. <https://doi.org/10.1016/j.socscimed.2017.10.016>.
- McCalla, Sandra, Howard L. Minkoff, Joseph Feldman, Isaac Delke, Martin Salwin, Gloria Valencia, and Leonard Glass. 1991. "The Biologic and Social Consequences of Perinatal Cocaine Use in an Inner-City Population: Results of an Anonymous Cross-Sectional Study." *American Journal of Obstetrics and Gynecology* 164 (2): 625–30.
[https://doi.org/10.1016/S0002-9378\(11\)80036-0](https://doi.org/10.1016/S0002-9378(11)80036-0).
- Meyer, Bruce D. 2002. "Labor Supply at the Extensive and Intensive Margins: The EITC, Welfare, and Hours Worked." *American Economic Review* 92 (2): 373–79.
<https://doi.org/10.1257/000282802320191642>.
- Meyer, Bruce D., and Dan T. Rosenbaum. 2000. "Making Single Mothers Work: Recent Tax and Welfare Policy and Its Effects." *National Tax Journal* 53 (4): 1027–61.
- . 2001. "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *The Quarterly Journal of Economics* 116 (3): 1063–1114.
- Phibbs, C S, D A Bateman, and R M Schwartz. 1991. "The Neonatal Costs of Maternal Cocaine Use." *JAMA* 266 (11): 1521–26.
- Singer, Lynn T., Robert Arendt, Sonia Minnes, Kathleen Farkas, Ann Salvator, H. Lester Kirchner, and Robert Kliegman. 2002. "Cognitive and Motor Outcomes of Cocaine-Exposed Infants." *JAMA* 287 (15): 1952–60.
- Strully, Kate W., David H. Rehkopf, and Ziming Xuan. 2010. "Effects of Prenatal Poverty on Infant Health: State Earned Income Tax Credits and Birth Weight." *American Sociological Review* 75 (4): 534–62.
- Tax Policy Center 2019 <https://www.taxpolicycenter.org/statistics/eitc-recipients> (last accessed August 4, 2019).

- U.S. Department of Health and Human Services. *Women and Smoking: A Report of the Surgeon General*. Accessed March 21, 2018.
http://www.cdc.gov/tobacco/data_statistics/sgr/2001/complete_report/.
- Vega, William A., Bohdan Kolody, Jimmy Hwang, and Amanda Noble. 1993. "Prevalence and Magnitude of Perinatal Substance Exposures in California." *New England Journal of Medicine* 329(12): 850-4.
- Zuckerman, Barry, Deborah A. Frank, Ralph Hingson, Hortensia Amaro, Suzette M. Levenson, Herbert Kayne, Steven Parker et al. 1989 "Effects of maternal marijuana and cocaine use on fetal growth." *New England Journal of Medicine* 320(12): 762-768.

Tables and Figures

Figure 1: Maximum Available Credit from the EITC by Parity in 1995 \$



Source: HMS (2015) via Tax Policy Center (2011)

Notes: The relative change in available tax credits from the EITC is \$401 from 1990 to 1993 (\$222+\$179=\$401) among women of parity 2 relative to parity 1; the relative change is \$179 from 1991-1993 and \$264 from 1994-1996 (\$578-\$314=\$264) all in 1995 dollars.

Figure 2: Event-study estimates of low birth weight of parity 2 or parity 2+ relative to parity 1 for all single women with no more than a high school degree, (1993=0).

Figure 2A: All Women Parity 2 v. 1

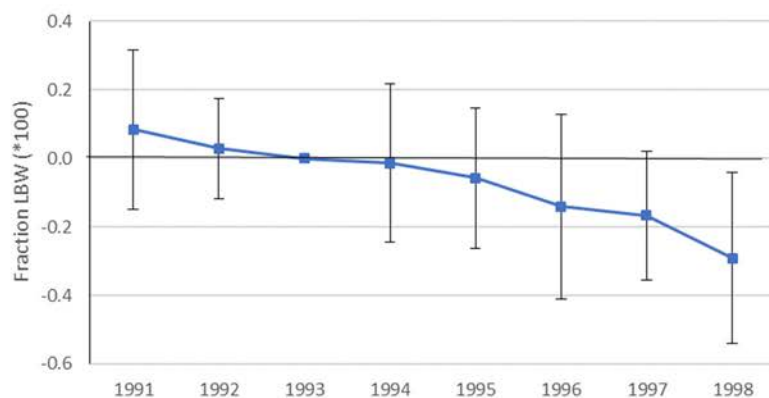


Figure 2B: All Women Parity 2 v. 1

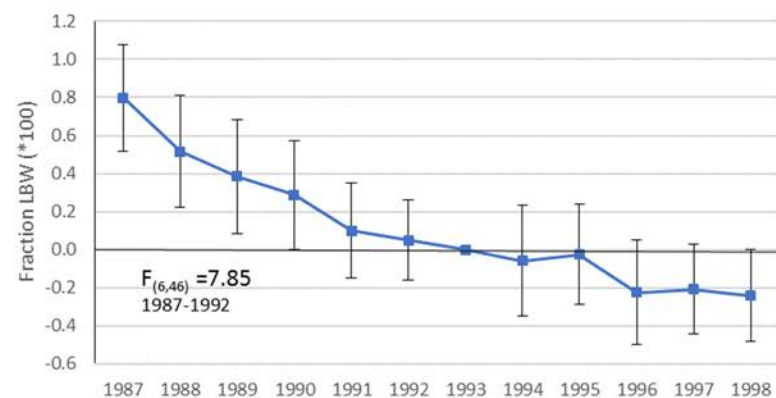


Figure 2C: All Women Parity 2+ v. 1

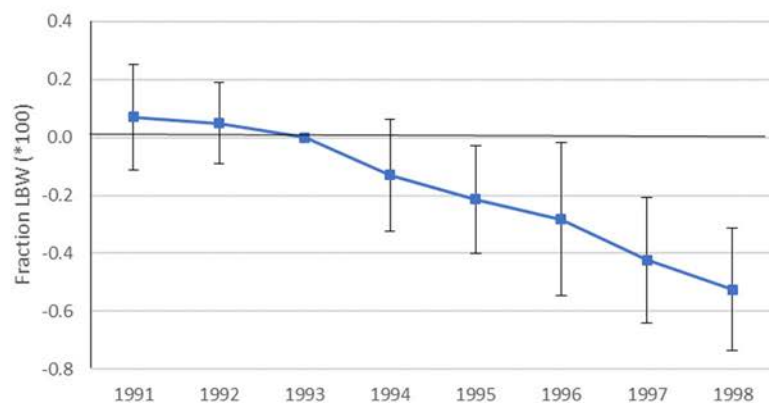


Figure 2D: All Women Parity 2+ v. 1

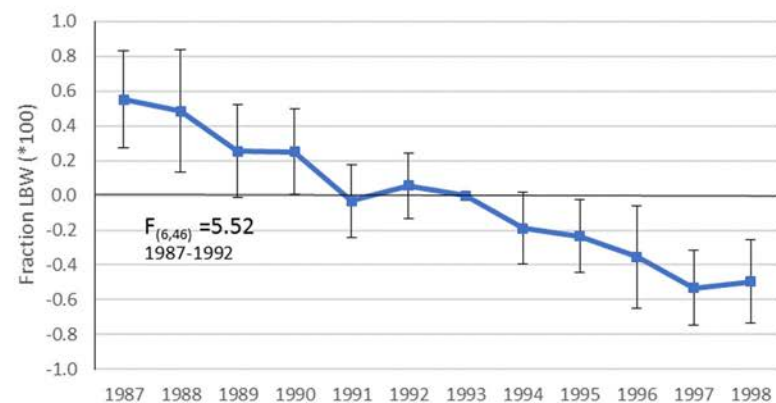


Table 1- Difference-in-differences Estimates of OBRA93 on Low Birth Weight Single Women with a High School Education or Less-Parity 2 to 1

Model	All (1)	Black (2)	White (3)	Hispanic (4)
Panel A: Pre-Period is 1991-1993 (HMS)				
Parity2 *After	-0.164** (0.072)	-0.310** (0.144)	-0.114* (0.072)	-0.060 (0.07)
Mean LBW	10.2	14.4	8.1	7.0
Observations	47,687	13,780	21,775	14,823
Panel B: Pre-Period is 1991-1993[^]				
Parity2 * After	-0.209*** (0.072)	-0.390*** (0.138)	-0.146* (0.075)	-0.161 (0.195)
Mean LBW	9.9	13.3	8.3	7.5
Observations	21,677	6,155	9,997	6,771
Panel C: Pre-Period is 1987-1990				
Parity2 * After	-0.564*** (0.076)	-0.920*** (0.147)	-0.352*** (0.070)	-0.288* (0.167)
Mean LBW	10.1	13.5	8.3	7.6
Observations	32,399	9,237	14,794	9,744
Panel D: Pre-Period is 1987-1990				
Parity2+ * After	-0.604*** (0.094)	-1.064*** (0.169)	-0.229*** (0.080)	-0.350*** (0.126)
Mean LBW	11.1	14.8	8.7	7.9
Observations	63,240	18,155	29,057	18,763

[^] Estimates in Panel A are from HMS Tables 2 and 3 from a specification that pools women of all parities. Estimates change only slightly if we estimate the same regressions but limit the sample to women of parity 1 and 2. In Panel B we use 1991-1993 as the pre-period, but we drop California, New York, Texas, and Washington because they did not report mother's schooling prior to 1989. In Panels C and D we use the same states as in Panel B but we use 1987-1990 as the pre-period.

Table 2- Difference-in-differences Estimates of OBRA93 on Low Birth Weight Single Women with a High School Education or Less (1991-1998)-Placebo Comparisons

Model	All (1, HMS [^])	Black (2)	White (3)	Hispanic (4)
Pane A: Parity 3+ v. 2 (HMS[^])				
Parity3+*After	-0.347*** (0.067)	-0.715*** (0.121)	-0.028 (0.073)	-0.121 (0.092)
Mean of dep var	10.7	14.9	8.2	6.8
Observations	35,467	10,273	16,247	10,951
Pane B: Parity 3 v. 2				
Parity3 * After	-0.225*** (0.062)	-0.476*** (0.135)	-0.046 (0.085)	-0.05 (0.073)
Mean of dep var	9.7	13.4	7.8	6.3
Observations	23,916	6,865	10,967	7,422
Pane C: Parity 3 v. 4				
Parity4 * After	-0.240*** (0.089)	-0.421*** (0.135)	-0.020 (0.108)	-0.124 (0.216)
Mean of dep var	11.1	15.2	8.6	6.8
Observations	22,021	6,326	10,381	6,625
Pane D: Parity 3 v. 4+				
Parity4+ * After	-0.261** (0.105)	-0.483*** (0.136)	0.062 (0.134)	-0.107 (0.142)
Mean of dep var	12.0	16.5	9.0	7.3
Observations	38,675	11,352	18,593	11,041

[^] All estimates in column 1 and Panel A were included in HMS' Table 2, 3, or Appendix Table 6 except parity 3 v 4+ . The estimates between this table and ? differ only slightly from their published results because we corrected HMS's coding error as they did in a response to us. We agree the coding error made no practical difference.

Table 3- Difference-in-differences Estimates of OBRA93 on Any Smoking for Single Women with a High School Education or Less (1991-1998)-Placebo Comparisons

Model	All (1 [^])	Black (2)	White (3)	Hispanic (4)
Panel A: Parity 3 vs 2				
Parity3*After	-0.935 (0.148)***	-1.837 (0.215)***	-0.307 (0.176)*	-0.627 (0.184)***
Mean of dep var	27.2	15.5	35.5	8.2
Observation	22,863	6,865	10,603	7,422
Panel B: Parity 4 v. 3				
Parity4*After	-0.494 (0.212)**	-1.233 (0.262)***	0.548 (0.287)*	-0.334 (0.350)
Mean of dep var	30.1	21.1	38.2	10.6
Observation	20,959	5,983	9,995	6,227
Panel C: Parity 4+ v. 3				
Parity4+*After	-0.514 (0.249)**	-1.18 (0.136)***	0.648 (0.247)**	-0.217 (0.353)
Mean of dep var	31.3	24.5	38.3	11.6
Observation	36,541	11,352	17,751	10,295

[^] All estimates in column (1) of table were included in HMS Appendix table 6 except 4+ to 3. The estimates between this table and table 6 differ only slightly because we corrected HMS's coding error by including a full set of interactions for the demographic fixed effects. The coding error made no practical difference.

Table 4- Difference-in-differences Estimates of OBRA93 on Low Birth Weight of Single Women with no more than a High School Education of Parity 3+ versus Parity 2

	Full Model	Limited Model	EITC+ Welfare Reform	Welfare Reform	EITC+ Unemployment Rate	Unemployment Rate	EITC + Medicaid	Medicaid
	(1, HMS)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All Women								
Parity3+ * After	-0.340*** (0.068)	-0.349*** (0.076)	-0.176** (0.085)		-0.087 (0.110)		-0.342** (-0.131)	
Parity3+*Policy			-0.308** (0.144)	-0.492*** (0.108)	0.142** (0.067)	0.178*** (0.038)	-0.001 (0.013)	-0.027*** (0.009)
Panel B: Black Women								
Parity3+ * After	-0.715 (0.121)***	-0.594 (0.113)***	-0.169 (0.171)		-0.067 (0.219)		-0.528*** (0.178)	
Parity3+*Policy			-0.977** (0.294)	-1.159*** (0.204)	0.354*** (0.120)	0.383** (0.160)	-0.023 (0.020)	-0.062*** (0.013)

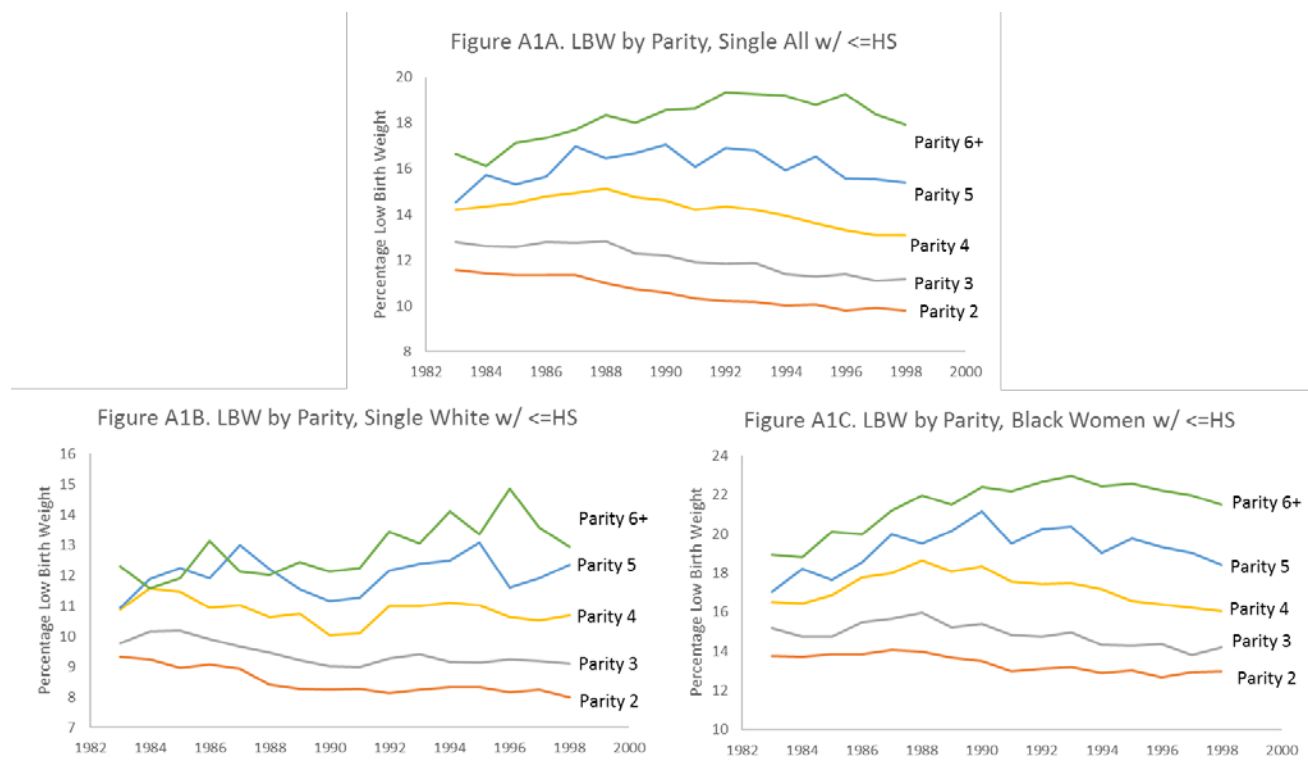
Figures show the estimated effect of the 1993 EITC expansion on women of parity 3 or higher relative to women of parity 2. Estimates in the full model are from equation (1) in the text and include all state and demographic fixed effects and the 3 state variables that vary by state and year. The limited model drops all state and demographic fixed effects as well as the 3 state variables and includes only parity and year fixed effects along with the interaction of the after period and the indicator of parity 3 or higher. In column (3) we add the annual proportion of women in the HMS sample exposed to welfare reform interacted with the indicator for parity 3+ to the full model. We drop the state measure of welfare reform but keep state-year variation in the other two state variables. In column (4) we drop the parity 3+*after interaction. We do the same for unemployment rate in columns (5) and (6) as well as the measure of Medicaid/SCHIP eligibility thresholds in columns (7) and (8). Standard errors in parentheses are clustered at the state level following HMS. There are 35,467 observations for all women and 10,273 observations for Black women.

Appendix A

AI. Trends in Low Birth Weight by Parity and Race

A key assumption in difference-in-difference designs is that trends in those exposed to an intervention are similar or roughly parallel to those unexposed in the periods leading up to the change. In the three panels below we show the rate of low birth weight by race and parity for all single unmarried women in Figure A1a, white women in Figure A1b, and Black women in Figure A1c. The horizontal axis shows what HMS refer to as the effective tax year which defines a woman's exposure to the EITC. What is apparent among all women and especially Black women is the disparity in the time-series pattern of low birth weight among women of parity 5 and 6+ relative to women of parity 2, 3 and 4. This becomes important in our placebo analysis in the text. Women of parity 3 and higher all received the same increase in the EITC after 1993. A natural placebo test is whether the EITC had a differential impact on low birth weight between women of parity 3 and 4 as well as women of parity 4+. A concern of including women of parity 5+ is the differential trends in low birth weight between them and women of parity 3. Thus, in all our placebo tests we compare women of parity 4 to parity 3 separately from women of parity 4+ and 4.

The other noticeable pattern is the curvilinear trajectory of the series with peaks among women of parity 2, 3 and 4 around 1988 followed by steady declines. These patterns become important in the DD analysis. They suggest that rates of low birth weight were declining well before the 1993 expansion among Black women of parities 2, 3 and 4. Although the scale of the vertical axis makes it difficult to discern small changes, there is no clear discontinuity in low birth weight after 1993 among women of parity 3 and 4 relative to parity 2. Second, the curvilinear pattern becomes relevant when HMS estimate the effect of the EITC expansions in 1986, 1990 and 1993 together. As we show in Appendix A section IV, estimates from that analysis are sensitive to the inclusion of trend terms interacted with parity.



All. Event-study plots of the placebo tests

As noted in the text, women of parity 3, 4 and 4+ all experienced the same increase in the EITC after 1993. Consequently, there should be no differential effect of the EITC on the rate of low birth weight between women of parity 4 or 4+ relative to parity 3. As we show in the text, Black women fail this placebo test. It is important, however, to insure that the failed placebo tests are not the result of differential trends in low birth weight prior to the expansion. In Figures A2 we estimate event-study regressions of the low birth weight among black women for HMS's study period of 1991-1998. Figures A2a, A2b and A2c compare black women of parity 3 versus 2, parity 4 versus 3 and parity 4+ versus 3, respectively. The pattern in each figure is the same: no differential pre-trends followed by similar declines in low birth weight for each comparison. The graphs demonstrate that the failed placebo tests between women of parity 4 and 4+ versus parity 3 are not the result of differential trends in low birth weight in the pre-period. We discuss potential confounders in Sections V of the text.

Figure A2: Event-study estimates of low birth of for all single Black women with no more than a high school degree of 3 v. 2, 4 v. 3 and 4+ v. 4, (1993=0)

Figure A2A: Black Women Parity 3 v. 2

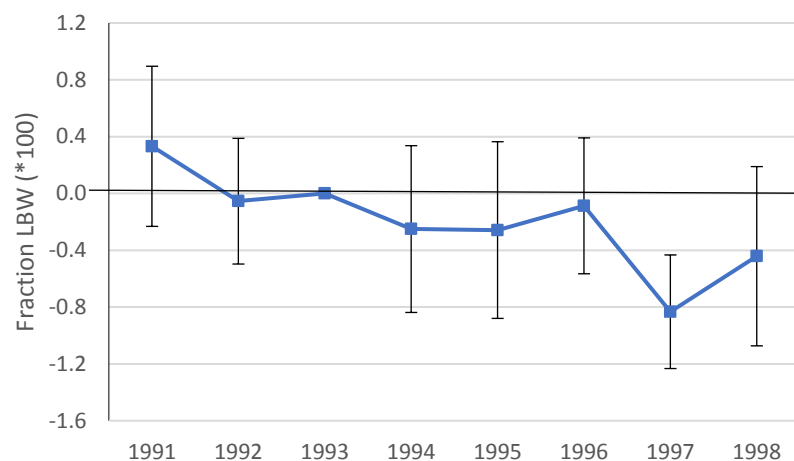


Figure A2B: Black Women Parity 4 v. 3

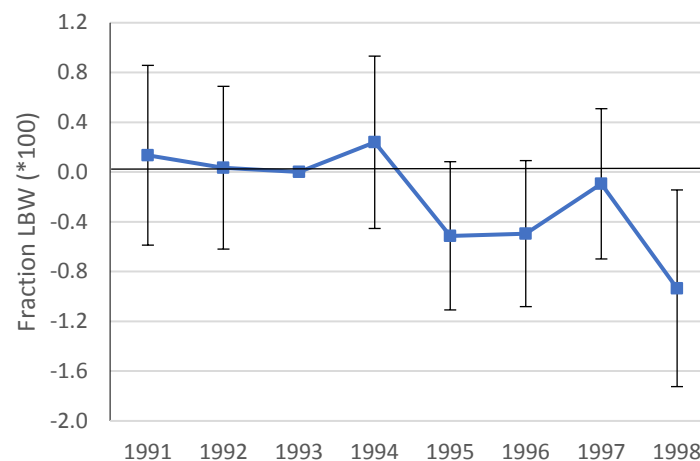
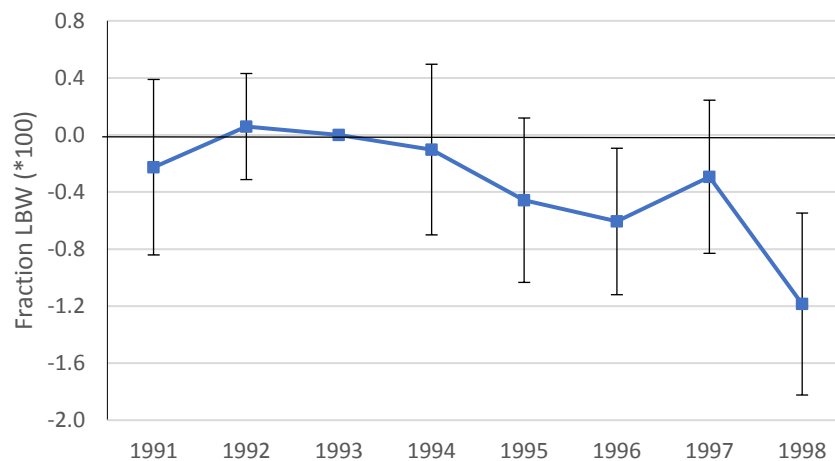


Figure A2C: Black Women Parity 4+ v. 3



AIII. Possible Mechanisms: Prenatal Care

We find prenatal care to be an implausible explanation. First, the association between prenatal care and birth outcomes in the public health literature is modest at best based on observational studies (Corman, Dave and Reichman, forthcoming). The randomized studies of prenatal or augmented prenatal care have reported almost uniformly, no association with improved infant health (Collaborative Group on Preterm Birth Prevention 1993; Goldenberg and Rouse 1998; Goldenberg and Culhane 2007; Alexander and Kotelchuck 2001; Klerman et al. 2001; Carroli et al. 2001; Corman Dave and Reichman, forthcoming). The multicenter RCT of preterm prevention is a particularly germane example (Collaborative Group on Preterm Birth Prevention 1993). The intervention targeted women at high risk for preterm birth. The treatment group received weekly examinations beginning in weeks 20-24 until delivery. Patients in the treatment group were also trained to recognize the signs of preterm labor. The intervention conferred no benefit despite a level of support that far exceeded routine prenatal care. In a more recent review of the determinants of preterm birth, the author wrote,

Strategies to prevent preterm birth have traditionally emphasized early prenatal care as providing an opportunity to identify and treat prematurity-related risk factors, but this approach has not reduced the incidence of preterm birth. Improved access to prenatal care is associated with lower rates of preterm birth, but the linkage is apparently related more to the high rates among women who receive no prenatal care than to the content of care received (Iams 2014, p. 256)

Even based on the questionable estimates of prenatal care effectiveness in the public health literature, the change in prenatal care associated with the EITC is inconsequential. For instance, in Table A1 below we replicate HMS's prenatal care results for all women of parity 3+ versus parity 2 (HMS, Table 7). HMS report that the EITC is associated with an increase of 0.652 percentage points in prenatal care initiation before the third trimester. This is less than a one-percent increase evaluated at the mean of 89.42 percent, a change so small as to be clinically irrelevant.

Lastly, trends in prenatal care vary by parity and the effect of prenatal care fails the placebo test. As we noted in the text, women of parity 3, 4 and 4+ all receive the same increase in the EITC after 1993. There should be, therefore, no differential change in prenatal care between women of parity 4 versus 3 or parity 4+ versus 3. Again, refer to Table A1. The EITC is associated with a 0.493 percentage point increase between all women of 4 versus parity 3 and a 0.659 percentage point increase among black women. These changes are small and not clinically important, but they point to differential trends in prenatal care by parity that are consistent with the differential trends in low birth weight and prenatal smoking discussed in the text. Also of note is the lack of an association between the EITC and prenatal care among white or Hispanic women of parity 3+ versus 2. This is consistent with the lack of an association between the EITC and low birth weight between white and Hispanic women of parity 4 versus 3.

Table A1- Difference-in-differences Estimates of OBRA93 on the Percent Births in Which Prenatal Care Was Initiated in the 2nd Trimester or Earlier for Single Women with a High School Education or Less by Parity and Race/Ethnicity

Model	All (1 [^])	Black (2)	White (3)	Hispanic (4)
Parity 3+ vs. 2				
Parity3 * After	0.651 (0.175)***	1.244 (0.029)***	0.119 (0.133)	0.161 (0.174)
Mean Dep Variable	89.42	87.48	90.73	89.13
Observations	35,141	10,132	16,174	10,830
Parity 3 vs. 2				
Parity3 * After	0.426 (0.111)***	0.937 (0.167)***	0.095 (0.102)	0.067 (0.0137)
Mean Dep Variable	91.11	89.99	91.8	90.16
Observations	23,717	6,775	10,928	7,347
Parity 4 vs. 3				
Parity4 * After	0.493 (0.103)***	0.659 (0.192)***	0.209 (0.138)	0.222 (0.227)
Mean Dep Variable	88.73	87.02	89.97	88.89
Observations	21,811	6,237	10,327	6,549
Parity 4+ vs. 3				
Parity4+ * After	0.504 (0.177)***	0.651 (0.306)**	0.194 (0.120)	0.075 (0.166)
Mean Dep Variable	87.08	84.72	88.91	87.85
Observations	38,231	11,177	18,457	10,906

[^] The dependent variable is the percent of births to women who initiated prenatal care prior to the third trimester as recorded on the birth certificates. All estimates use data posted by HMS. Each estimate is from a separate regression as described by Equation (1) in the text. The standard errors have been adjusted for clustering at the state level following HMS.

References

- Alexander, Greg R., and Milton Kotelchuck. 2001. "Assessing the Role and Effectiveness of Prenatal Care: History, Challenges, and Directions for Future Research." *Public Health Reports* (1974-) 116 (4): 306–16.
- Corman, Hope, Dhaval Dave and Nancy Reichman. Forthcoming. "Effect of Prenatal Care on Birth Outcomes: Reconciling a Messy Literature" *Southern Economic Journal*.
- Carroli, Guillermo, José Villar, Gilda Piaggio, Dina Khan-Neelofur, Metin Gülmezoglu, Miranda Mugford, Pisake Lumbiganon, Ubaldo Farnot, and Per Bersgjø. 2001. "WHO Systematic Review of Randomised Controlled Trials of Routine Antenatal Care." *The Lancet* 357 (9268): 1565–70. [https://doi.org/10.1016/S0140-6736\(00\)04723-1](https://doi.org/10.1016/S0140-6736(00)04723-1).
- Collaborative Group on Preterm Birth. 1993. Multicenter randomized controlled trial of a preterm birth prevention program. *American Journal of Obstetrics and Gynecology* 169(2 pt):352-366.
- Goldenberg, Robert L., and Jennifer F. Culhane. 2007. "Low Birth Weight in the United States." *The American Journal of Clinical Nutrition* 85 (2): 584S–590S. <https://doi.org/10.1093/ajcn/85.2.584S>.
- Goldenberg, Robert L., and Dwight J. Rouse. 1998. "Prevention of Premature Birth." *New England Journal of Medicine* 339 (5): 313–20. <https://doi.org/10.1056/NEJM199807303390506>.
- Iams, Jay D. "Prevention of Preterm Parturition." Edited by Caren G. Solomon. *New England Journal of Medicine* 370, no. 3 (January 16, 2014): 254–61. <https://doi.org/10.1056/NEJMcp1103640>.
- Klerman, L. V., S. L. Ramey, R. L. Goldenberg, S. Marbury, J. Hou, and S. P. Cliver. 2001. "A Randomized Trial of Augmented Prenatal Care for Multiple-Risk, Medicaid-Eligible African American Women." *American Journal of Public Health* 91 (1): 105–11.

AIV. The EITCs of 1986, 1990 and 1993

As noted in the text, HMS expand their analysis of the EITC and low birth weight by combining the EITC expansions in 1986 and 1990 with that of 1993.¹ To evaluate the impact of all three expansions, HMS again aggregate individual-level birth certificate data to cells defined by state, year, parity, maternal education, race, ethnicity and age from 1983 to 1998 (HMS, Appendix B). They limit the sample to single, unmarried women who gave birth. To account for the multiple expansions HMS estimate the following two-way fixed effects specification:

$$(2) Y_{pjst} = \alpha + \delta Maxcredit_{pt} + \pi X_{st} + \rho_p + \varphi_j \varphi_k + \delta_t + \tau_s + \epsilon_{pjst}$$

Y_{pjst} is the rate of low birth by parity (p), demographic groups (j), state (s) and year (t). X is a set of state policies: Medicaid/SCHIP, welfare reform and the state unemployment rate. $Maxcredit_{pt}$ is the maximum tax credit available to eligible filers that varies by parity and year. HMS include an additional term ($p_p * T$) to control for linear trends by parity. We replicate HMS's analysis in the top Panel of Table A2 below. The coefficient for all women indicates that a \$1000 increase in the maximum available credit is associated with -0.304 percentage point decline in the rate of low birth weight. The effects for white women are substantially smaller [-0.117, column (4)], while those for black women are much larger [-0.518, column (7)].

The estimates in Table A2 are also sensitive to the inclusion of trend terms by parity. Inclusion of a linear trend term almost triples the estimated effect for black women from -0.518 to -1.357 percentage points (columns 7 and 8), whereas a quadratic trend in parity eliminates any effect of the EITC on low birth weight (column 9). Trend terms interacted with parity may over fit the data, but as we show in Appendix A, Figure A1, the time-series pattern in low birth weight is clearly curvilinear especially

¹ The Tax Reform Act, 1986; Omnibus Budget Reconciliation Act, 1990; and the Omnibus Budget Reconciliation Act, 1993 (Nichols and Rothstein 2016).

among black women. At a minimum, the effect of the EITC expansions on low birth weight is observationally equivalent to nonlinear trends in low birth weight by parity unrelated to the EITC.

In the lower panel of Table A2 we re-estimate the same model but limit the sample to the effective years 1983-1993. In this subsample there is no association between the $Maxcredit_{pt}$ and low birth weight for all women and all black women. Inclusion of a linear trend in parity eliminates any effect for white women. The lack of an association between the EITC expansions in 1986 and 1990 is not surprising given that take-up and the magnitude of the available tax credits were substantially less in these early expansions. The point, however, is that the estimated effect from the extended analysis is driven by the 1993 expansion, which we have argued lacks convincing evidence of a causal effect.

Table A2- Maximum Credit Estimates of EITC on Low Birth Weight, Single Women with a High School Education by Race

Model	All			White			Black		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: 1983-1998									
Maximum Credit (\$1,000 of 95\$)	- 0.304*** (0.066)	-0.772*** (0.128)	-0.075 (0.150)	-0.117** (0.053)	-0.075 (0.107)	0.005 (0.159)	-0.518*** (0.114)	-1.357*** (0.179)	-0.0214 (0.220)
Parity X linear time		X	X		X	X		X	X
Parity X quadratic time			X			X			X
P-value, trends terms		0.000	0.000		0.000	0.000		0.000	0.000
Mean LBW	11.21	11.21	11.21	8.81	8.81	8.81	14.76	14.76	14.76
Observation	81,782	81,782	81,782	37,335	37,335	37,335	23,746	23,746	23,746
Panel B: 1983-1993									
Maximum Credit (\$1,000 of 95\$)	-0.087 (0.147)	-0.330 (0.360)	0.527* (0.291)	-0.467*** (0.119)	0.004 (0.362)	0.072 (0.361)	0.147 (0.284)	-0.359 (0.552)	1.14 (0.478)**
Parity X linear time		X	X		X	X		X	X
Parity X quadratic time			X			X			X
P-Value, trend terms		0.000			0.000	0.000		0.000	0.000
Mean LBW	11.43	11.43	11.43	8.83	8.83	8.83	14.88	14.88	14.88
Observation	55,003	55,003	81,782	24,805	24,805	24,805	16,220	16,220	16,220

AV. Crack Cocaine as a possible Confound

In this Section of the Appendix we use birth certificate data from the New York City Department of Health and Mental Hygiene to show trends in low birth weight, prenatal exposure to narcotics by race and parity. From 1980 to 1987 there was an indication for prenatal exposure to narcotics as a medical risk factor on the New York City birth certificate. Beginning in 1988, the indication was refined to include separate codes for heroin, cocaine and marijuana. We tabulate the percent of births with an indication of narcotic use during pregnancy from 1980-1987 and splice it with the percent of births with an indication of cocaine and heroin from 1988 to 2000 by race and parity.² In Appendix Figure A3 Panel A we show the percent of low birth weight births and the percent of births exposed prenatally to illicit drugs. The data pertain to black women only with separate series for women of parity 2 and parity 3+. The data demonstrate that the change in prenatal drug use rose faster among black women of parity 3+ relative to parity 2 beginning around 1985 as did the rate of low birth weight. In Panel B we mimic HMS's event-study by showing the differential change in low birth weight and drug use normalized to 1993. We show the same two panel for white women in Appendix Figure A4. The data for white women are noisier with a less obvious distinction in low birth weight and drug use by parity.

Next, we overlay the differential rate of prenatal drug use between black women of parity 3+ relative to parity 2 with the NYC homicide rate of black males 15 to 24 (Appendix Figure A5). As noted in the text (Section III) numerous social scientists have linked the rise in homicide among young black males to growth in crack cocaine markets in urban areas. The New York City homicide rate among young black males closely tracks the rise in prenatal exposure to cocaine as well the rate of low birth weight among black women. To link the patterns in New York City to national data, we use plot the

² Screens for exposure to cocaine as reported on birth certificates are based on self-reports and may include a physician's indication based on the medical chart. There is little doubt that the true prevalence is underreported (Behnke et al. 2013).

differential rate of low birth weight of black women with a high school degree or less between women of parity 3+ and parity 2 using national data from HMS and the national rate of homicide among black males 15-24 years of age (Appendix Figure A6, Panel A). In Panel B we show the same figure for white women. All series are normalized to 1993 so as to make them comparable to HMS's event study figures. The point is twofold: First, low birth weight, drug use and homicide rates begin to rise after 1985, peak around 1991 and decline thereafter. Second, to the extent that the rise in homicide rates reflects the spread of crack-cocaine markets, they are consistent with the clinical literature, which reported greater exposure among black women of higher relative to lower parity (Vega et al. 1993).

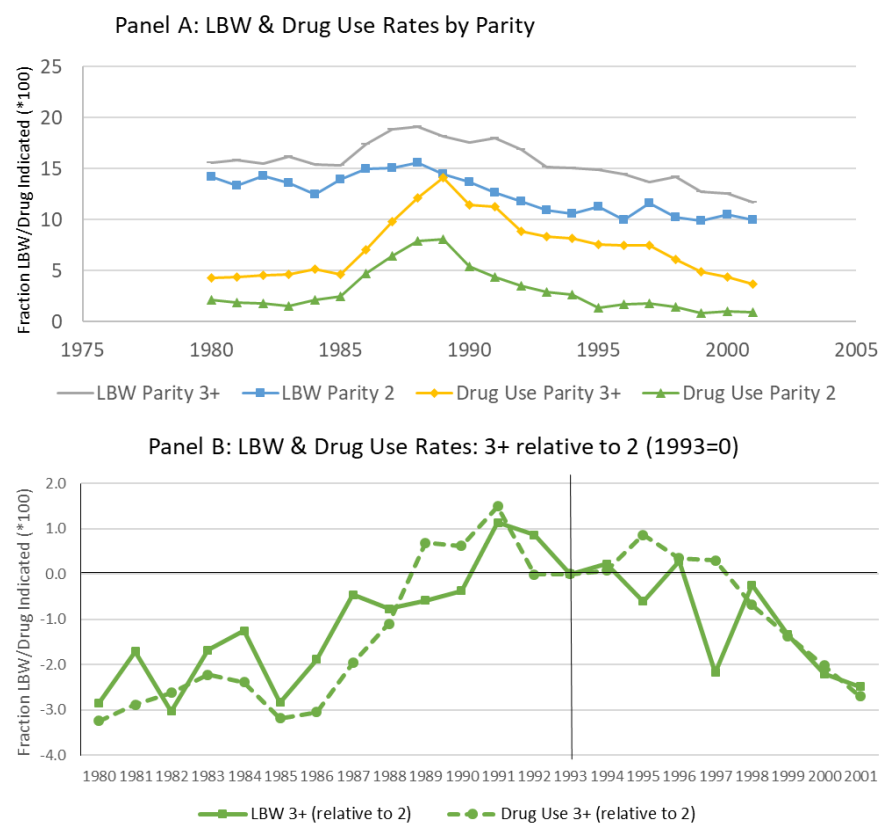
There are no comparable data of prenatal exposure to cocaine by race and parity from other parts of the country. However, as noted in footnote 16 of the text, we cite studies of prenatal exposure to cocaine and birth outcomes from hospitals and clinics in Boston, Chicago, Pittsburgh, and Florida. The unprecedented prevalence study in scale and accuracy by Vega et al. (1993) confirms that the differential exposure to cocaine by race was not limited to New York City was evident across the entire state of California.

Re-estimating the DDs with Homicide Rates

In this section we use the national homicide rate among black males 15-24 as a proxy for the spread of crack-cocaine markets in the 1980s and 1990s. We add the homicide rate interacted with the women of parity 3+ to HMS's DD specification of low birth weight between women of parity 3+ and 2. The results in Appendix Table A3 (Panel A) show that there is no longer any association between 1993 EITC expansion and low birth weight among black women of parity 3+ relative to parity 2. The null results for white and Hispanic women reported by HMS remain as such. We then estimate comparisons between women of parity 4 and 4+ relative to parity 3. Recall that this placebo test found substantial differences between these two groups of black women when there should have been none. As we show

in Panels B and C of Appendix Table A3, inclusion of the black homicide rate among young black males interacted with parity eliminates this association as well. Put differently, inclusion of the homicide rate interacted with parity in the DD regressions appears to eliminate the omitted variable bias suggested by failed placebo tests reported in the Table 2 of the text. We cannot draw any causal conclusions from this exercise. Nevertheless, the patterns we have detailed by race and parity are consistent with the possible confounding of the EITC and low birth weight in HMS's analysis.

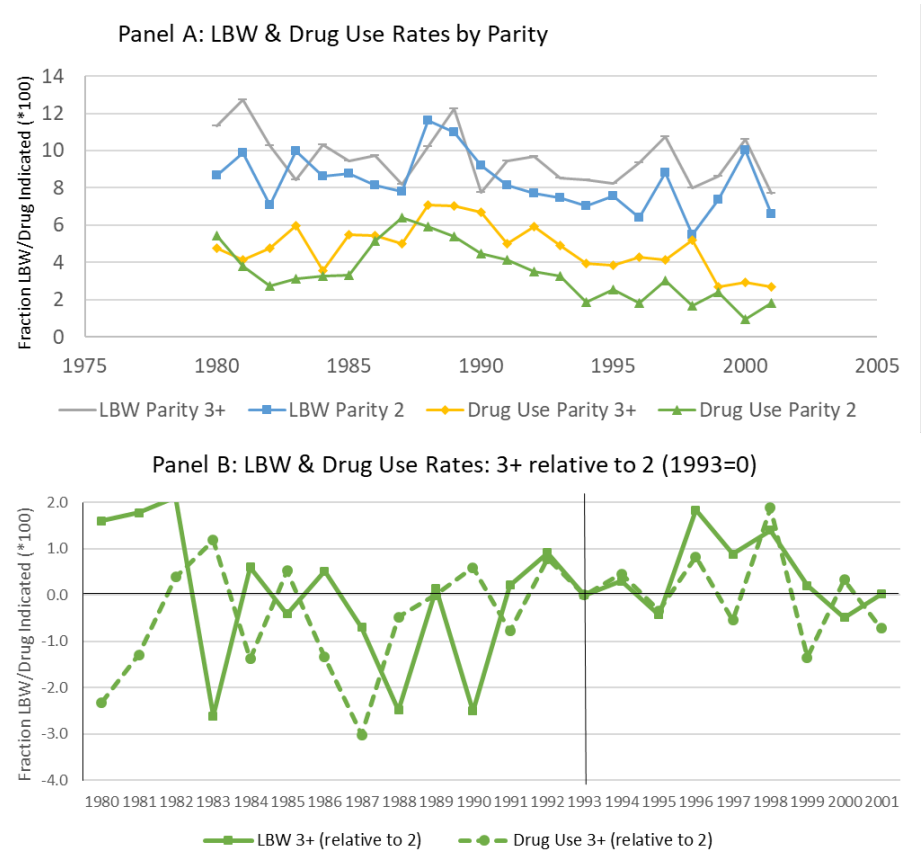
Figure A3: Percent of Births Exposed Prenatally to Narcotics and Cocaine (Drug Use) and Percent Low Birth Weight (LBW) Among Single Black Women with at most a High School Diploma in New York City by Year and Parity



Source: Authors' tabulations of NYC Birth Certificates (1980-2001)

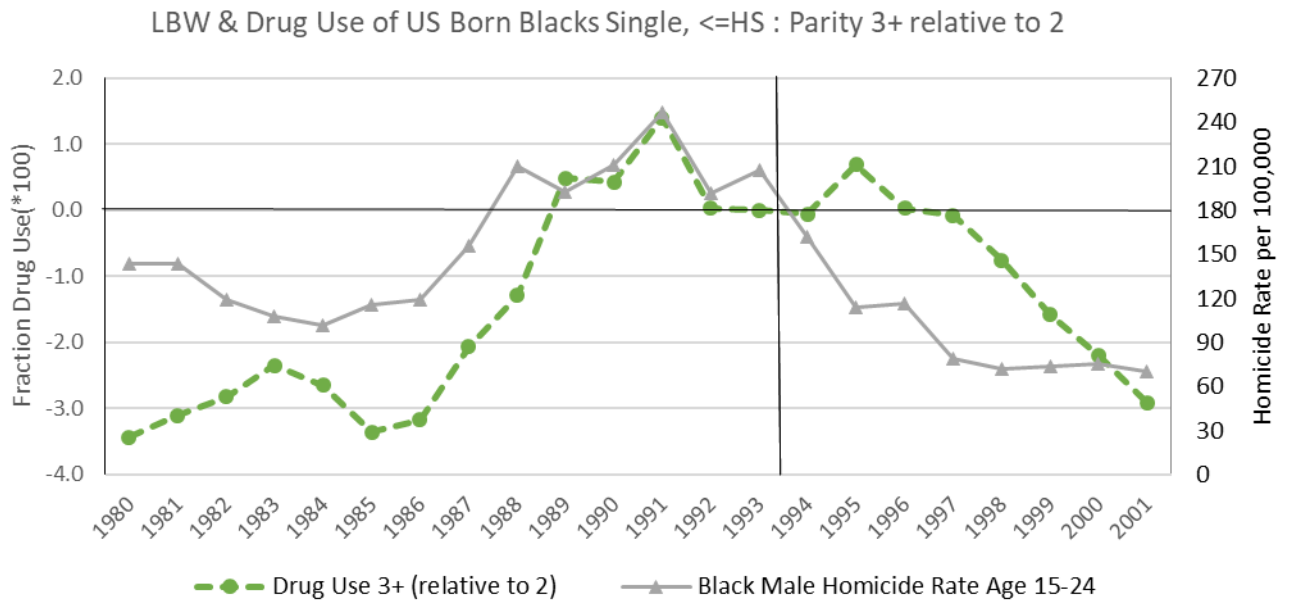
Notes: Panel A shows the absolute rate of low birth weight and prenatal use of narcotics from 1980-1987 and heroin and cocaine from 1988-1998 based on indications on the birth certificate. We refer to this as prenatal drug use. Panel B shows the same for parity 3+ relative to 2 using 1993 as the reference year.

Figure A4: Percent of Births Exposed Prenatally to Narcotics and Cocaine (Drug Use) and Percent Low Birth Weight (LBW) Among Single White Women with at most a High School Diploma in New York City by Year and Parity



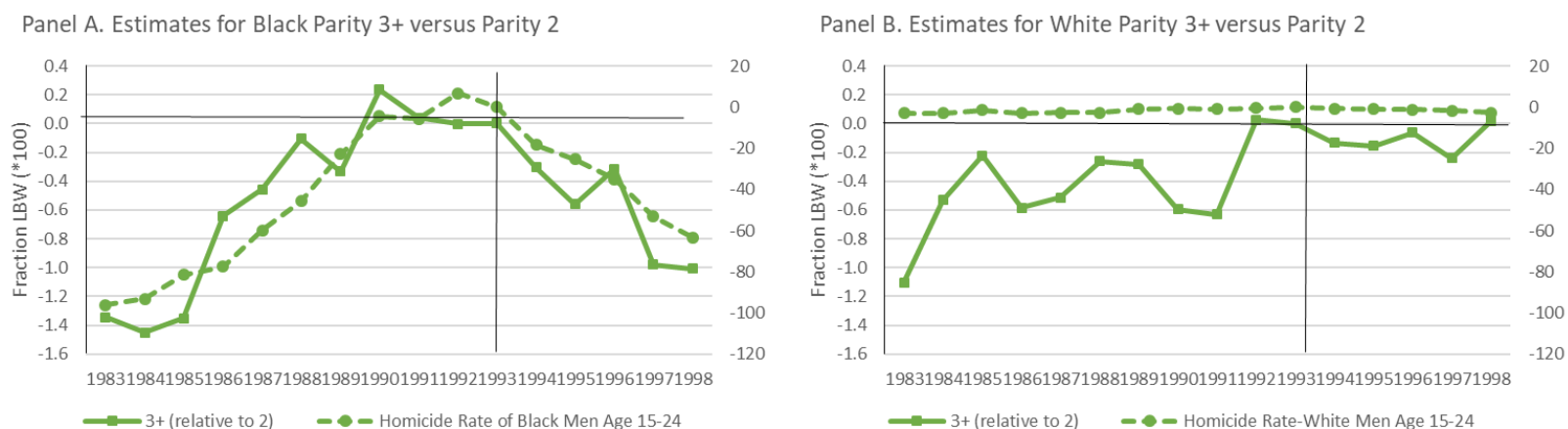
Source: Authors’ tabulations of NYC Birth Certificates (1980-2001). See Note to Figure 3.

Figure A5: Homicide Rates of Black Males 15-24 Years of Age Separately in New York City with the Percent of Births Exposed Prenatally to Narcotics and Cocaine (Drug Use, (1993=0)) Among Single Black Women of Parity 3+ Relative to Parity 2 with at most a High School Diploma in New York City



Source: Authors' tabulation of NYC birth certificates and Multiple Cause of Death Files (1980-1988) and Compressed Mortality Files (1989-2001). We thank Tim Moore for data on homicides (see Evans, Garthwaite and Moore 2018).

Figure A6: Event-Time Estimates of Low Birth Weight of Women of Parity 3+ Relative to Parity 2 Among Single Women with at most a High School Diploma Overlaid with National and Homicide Rates for Males Ages 15-24 for Black Women (Panel A) and White Women (Panel B), (1993=0)



Source: HMS (2015) and Multiple Cause of Death Files (1980-1988) and Compressed Mortality Files (1989-2001). We thank Tim Moore for data on homicides (see Evans, Garthwaite and Moore 2018).

Notes: Both panels contrast women of parity 3+ versus parity 2. They exclude California, New York, Texas and Washington because those states are missing education in some years preceding effective year 1991

Table A3- Difference-in-differences Estimates of OBRA93 on Low Birth Weight Single Women with a High School Education or Less Controlling for the Black Homicide Rate, Ages 15-24

Model	All (1 [^])	Black (2)	White (3)	Hispanic (4)
Panel A: Parity 3+ v. 2				
Parity3 * After	-0.080 (0.109)	0.013 (0.218)	-0.032 (0.146)	-0.145 (0.226)
Mean of dep var	11.3	15.4	9.9	8.7
Observations	35,467	10,273	16,247	10,951
Panel B: Parity 4 v. 3				
Parity4 * After	0.109 (0.224)	0.215 (0.293)	0.149 (0.379)	0.147 (0.672)
Mean of dep var	11.1	15.2	8.6	6.8
Observations	22,021	6,326	10,381	6,625
Panel C: Parity 4+ v. 3				
Parity4 * After	0.250 (0.118)	0.295 (0.259)	0.370 (0.291)	0.160 (0.426)
Mean of dep var	11.8	16.1	10.3	8.1
Observations	23,237	6,759	10,689	7,422

[^] All estimates are from Equation (1) in the text. Each include the annual national homicide rate for Black males 15-24 interacted with Parity 3+ (Panel A) or parity 4+ (Panel B). Standard errors are clustered at the state level following HMS.

Appendix B

In this Appendix we provide a more detailed review of the published articles on the effect of state EITCs on birth outcomes. The advantage of analyzing state EITCs, as opposed to changes in the federal EITC, is the temporal and spatial variation in the availability of tax credits. Conceptually, the timing and geographic variation of tax credit expansions should lessen vulnerability to confounding relative to the federal EITC, which affects all states equally at a point in time. The state EITCs, however, are approximately one-fifth to one-tenth the magnitude of the federal EITC. Thus, their effects should similarly be less than those estimates obtained by the federal EITC. Second, state EITCs are also vulnerable to policy endogeneity in that the states that expanded their EITC are clearly not a random sample of states, but tend to be more generous along an array of social policies. Third, none of the state studies exploits the difference in available tax credits between women with 2 or more previous live births relative to women with one. This has been a major identification strategy in studies of labor supply and infant health with the federal EITC. Such comparisons would be a natural way to better control for within-state trends and would also provide a convenient placebo test between women with 2 versus 3 previous live births all of whom are exposed to the same tax credit. Lastly, several studies fail to correct the standard errors for clustering at the state level. As Bertrand, Duflo and Mullainathan (*Quarterly Journal of Economics*, 2004) demonstrate, canonical difference-in-differences analyses of state policies will greatly underestimate the standard errors if not adjusted for within-state autocorrelation.

Markowitz, Sara, Kelli A. Komro, Melvin D. Livingston, Otto Lenhart, and Alexander C. Wagenaar. "Effects of State-Level Earned Income Tax Credit Laws in the U.S. on Maternal Health Behaviors and Infant Health Outcomes." *Social Science & Medicine* 194 (2017): 67–75. <https://doi.org/10.1016/j.socscimed.2017.10.016>.

Komro, Kelli A., Sara Markowitz, Melvin D. Livingston, Otto Lenhart, and Alexander C. Wagenaar. "Effects of State-Level Earned Income Tax Credit Laws on Birth Outcomes by Race and Ethnicity." *Health Equity*. Accessed July 23, 2019. <https://www.liebertpub.com/doi/10.1089/heq.2018.0061>.

The authors use state changes in the EITC from 1994-2013 to assess the effect of an income transfer on maternal health behaviors and birth outcomes. There were 80 changes in state EITC programs during the study period that the authors categorize into 5 mutually exclusive categories:

- 1) no state EITC;
- 2) states with a non-refundable EITC with payments less than 10 % of the federal EITC;
- 3) states with a refundable EITC with payments less than 10 % of the federal EITC;
- 4) states with a non-refundable EITC with payments 10% or more of the federal EITC;
- 5) states with a refundable EITC with payments 10% or more of the federal EITC.

The authors use the census of U.S birth certificates and limit the data to mothers with a high school degree or less ($n > 30,000,000$). They find that mothers in states with a non-refundable EITC and with payments less than 10% of the federal EITC have infants 9.4 grams heavier ($p < .05$) and infants that are 0.3 percentage points less likely to be low birth weight ($p < .01$) than states with no EITC. The most generous states have infants 27 grams heavier and an 0.8 percentage point decrease in low birth weight relative to states with no EITC.

There are two key features of the federal and state EITC. First, the maximum eligible benefit varies by family size; and second, although the federal EITC is refundable, the EITC in some states is not refundable. The distinction between a refundable and non-refundable tax credit is significant. The federal EITC is refundable meaning if the family does not have a federal income tax liability, then the full amount of the tax credit is transferred to family as income. Ninety-six percent of families eligible for the federal EITC received tax refunds. Virginia, for example has a non-refundable state EITC that pays 20 percent of the federal EITC. Although 20 percent is relatively generous, many low-income earners have minimal state income tax liability and even if they did, they would not receive any income transfer in February or March, but would instead pay less in state income tax.

Variation in benefits by parity and refundability are useful for assessing the magnitude of the effects reported by Markowitz et al. (2017). Using 2004, the midpoint of their study period, the maximum federal benefit for a mother having her first birth was \$390, \$2,604 if having a second birth and \$4,300 if having a third or higher-order birth (all unadjusted for inflation). Thus, the maximum state benefit in 2004 in states that offer refunds less than 10% of the federal EITC would be at most \$39 for first-time mothers, \$260 for mothers having a second child and \$430 dollars for mothers having a third or higher order birth. Markowitz et al. find that low EITC states with non-refundable credits increase the birth weight of infants to single moms by 11 grams. This seems large when compared with Hoynes, Miller and Simon (2015, HMS), who look at the impact of the federal EITC, a benefit roughly 5-10 times greater, and find only a 9-gram increase (HMS, Table 6). Similarly, Markowitz et al. find that states with refundable EITCs greater than 10% of federal EITC, a benefit of roughly \$70, increase birth weights by 14 grams for first-time moms relative to states with no EITC ($p < .01$). Our concern is that trends within the states grouped by the generosity of the EITC may be confounding their effects. For example, consider the states with the most generous state EITCs. Women having a second or higher order birth have infants only 3.4 grams heavier than women having a first birth. The difference, 3.4 grams, is clinically inconsequential (Markowitz et al., Table 3, column 5, Panels C and D). And yet, the available tax credits to those of higher order births are an order of magnitude greater than those to women having a first birth. In fact, if we compare birthweight, low birthweight and gestational age between women having a second or higher order birth with women having a first birth *within* each level of EITC generosity, we would conclude that the state EITCs had no meaningful impact on infant health. This form of placebo test suggests confounding from other factors.

Policy endogeneity is another source of possible confounding. The states with the most generous EITCs by 2013 are Colorado, Connecticut, Illinois, Kansas, Massachusetts, Minnesota, Nebraska, New Jersey, New Mexico, New York, Vermont, and Washington (Markowitz et al. Figure 1). Except for Kansas and Nebraska, these states tend to be more liberal politically and based on AFDC maximum benefits, to have more generous social welfare policies (Kleven 2019). One suggestion would be for the authors to add an event-study design. If policy endogeneity is an issue, there may be improvement in infant health in the years leading up to the state EITCs. This could explain the differences in birth outcomes across levels of state generosity while showing little effect across parity within states of the same generosity.

Lastly, Markowitz et al. find that state EITCs have the same effect on the birth outcomes of both married and unmarried women. This is unexpected for two reasons. First, the federal EITC has had no effect on the labor supply of married women (Hoynes and Eissa 2004). If labor force participation is one of the mechanisms that improves self-esteem and possibly health, then the lack of any labor supply effects eliminates this pathway. Second, the EITC tax credits to families with two spouses working tend to be smaller than the credits to single women because more married couples have income that places them on the phase-out portion of the EITC (Hoynes and Eissa 2004).

The paper by Komro et al. (2019) has the same set of authors as Markowitz et al. (2017) and uses the same data and design. The focus is on the effect of state EITCs on birth outcomes by race and ethnicity. In this shorter paper, the authors do not present results by parity within groups of state EITC generosity. Thus, unlike with Markowitz et al. (2017) we cannot screen for possible confounding by looking at the effect of state EITCs on birth outcomes across parity with categories of state generosity.

Hamad, Rita, and David H. Rehkopf. "Poverty, Pregnancy, and Birth Outcomes: A Study of the Earned Income Tax Credit." *Paediatric and Perinatal Epidemiology* 29, no. 5 (September 2015): 444–52. <https://doi.org/10.1111/ppe.12211>.

Hamad and Rehkopf use a panel of women from the National Longitudinal Survey of Youth (NLSY) to analyze the separate effects of family income and state and federal EITC on children born between 1986 and 2000. They find that \$1000 in additional income increase birth weight by a clinically irrelevant 3.7 grams ($p < .10$) but that the \$1000 in EITC payments increases birth weight by 65 grams. The authors acknowledge that income is likely endogenous and they use the EITC payments as an instrument for household income. In the first stage they find that the EITC increases post-tax income by \$2,727 but the increase has no effect on birth weight.

We have several concerns with the specification and the estimation of standard errors in the study. For instance, the comparison group is unclear. They include married and unmarried women with family incomes of \$100,000 or less. They estimate EITC eligibility based on

income and the number of children and use the estimated EITC payment as the “treatment.” Families with zero payments are presumably the comparison group. They also include the mother’s hours or work and the spouse’s income if married, both of which are endogenous. Similarly, they control for measures of depression, self-esteem and locus of control and physical health, all of which have been used as outcomes in previous analyses of the EITC (Evans and Garthwaite 2014,). Another concern is that sample sizes are small ($n=3,938$). If the number of births are evenly distributed across the 14 years and 50 states, then there are 5.5 births per state/year cell ($3,938/700$). Third, the policy intervention is at the state-year level and thus the standard errors are likely underestimated because the true degrees of freedom for the EITC is 51 (50 states plus DC) and not the number of households. The IV results are also unexpected. The IV coefficient is the ratio of two covariances: $\text{Cov}(Y,Z)/\text{Cov}(X,Z)$ where Y is the outcome, Z the instrument and X the endogenous variable. They report a statistically significant reduced form $\{\text{cov}(Y,Z)\}$ and a strong first stage $\{\text{cov}(X,Z)\}$, but a statistically insignificant IV. This is an odd result and we suspect they have underestimated the standard errors in the reduced form (health on EITC payment). This means there is likely no statistically significant effect of the EITC on birthweight. Lastly they do not include state fixed effects, which would eliminate time-invariant differences between states, a standard practice in analyses of the EITC (Strully, Rehkopf and Xuan 2010; Markowitz et al. 2017). State fixed effects would be taxing on a model with relatively few observations per state, but they could have included likely state-level confounders such as welfare reform, the Medicaid eligibility expansions and state unemployment or they could have used a first difference model. The point is that during their study period there were profound changes in welfare reform, Medicaid/SCHIP and the macroeconomy. Absent these state-level controls could lead to confounding.

Strully, Kate W., David H. Rehkopf, and Ziming Xuan. “Effects of Prenatal Poverty on Infant Health: State Earned Income Tax Credits and Birth Weight.” *American Sociological Review* 75, no. 4 (2010): 534–62.

The authors use the census of birth certificates from 1980 to 2002 to test whether state EITCs affect infant health. They use two-way fixed effects models with a dummy variable that is one if the state implemented an EITC and zero otherwise. They include women with at least one previous life birth because the state tax credits for women with no children were non-existent until 1993 and never large. The authors find that state EITCs increase birth weight by 16 grams and lowers maternal smoking by 5 percent.

We are skeptical of these results along several dimensions. First only 8.3 percent of births to women with a high school degree or less were exposed to a state EITC from 1980-2002. The benefits amounts are relatively small, approximately 15 percent of the federal benefits, and yet the birth weight effects are double those reported by HMS (2015) in their analysis of the much larger federal EITC. The authors do not exploit the magnitude of state EITC and whether it is refundable or non-refundable. The range is substantial. Eleven of the 16 states in their sample began with EITC that were 10 percent or less of the federal EITC. Seven of those 11 states

subsequently raised the state EITC but the authors do not exploit these increases either. They also make no distinction between the much smaller tax credits available to women with one previous live birth as compared to women with more two or more previous live births that were part of the 1993 federal EITC expansion. This has been a major source of identifying variation in many studies of labor supply and health (HMS and Evans Garthwaite 2014). Similarly, they find state EITC increase the birth weight of women with no previous live births, a group whose benefits are so small (~\$40 on average) that any association between state EITCs and the birth weight of first births seems implausible. Lastly, they do not correct their standard errors for clustering at the state level, which would reduce the effective degrees of freedom from over 8 million to 51 (the number of states and DC). As an example of how much the standard errors are likely underestimated, the authors report that the variable, State WIC participation, lowers birth weight by a statistically significant 0.123 grams ($p < .001$). In other words, a 1 percentage point increase in state WIC participation would increase birth weight by .12 grams (.0036 of mean birthweight). We find it unlikely there is sufficient statistical power at the *state level* to detect an effect that is one-tenth of one gram.

Bruckner, Tim A., David H. Rehkopf, and Ralph A. Catalano. "Income Gains and Very Low-Weight Birth among Low-Income Black Mothers in California." *Biodemography & Social Biology* 59, no. 2 (July 3, 2013): 141–56.
<https://doi.org/10.1080/19485565.2013.833802>.

The authors analyze the effect of the federal EITC expansions from 1991 to 1996 on the rate of very low birth weight (VLBW) births among black women in California from 1989 to 1997. They use monthly data and time-series methods and find that the EITC was associated with a 37 percent *increase* in the odds of a VLBW ($n=108$). They test the robustness of this finding by estimating the odds of a VLBW birth with individual birth certificates from California over the same time period. Estimates from these regressions also find that VLBW increases by 31 percent among women exposed to the EITC.

The results are likely confounded by other factors. First, 99 percent of all VLBW births are preterm and the etiology of preterm birth is not well known. Hundreds of clinical trials have been unable to demonstrate interventions that can prevent preterm birth (Institute of Medicine 2007). Second, the authors provide no evidence of a mechanism. A similar time-series of prenatal smoking would have been useful. Third, the authors can only speculate why an increase in income would worsen infant health. Fourth, VLBW is a low prevalence outcome and we wonder why the authors did not analyze outcomes of greater frequency that could provide additional evidence of a negative effect. Lastly, as the authors acknowledge, with only one-state, their estimates are vulnerable to any time-varying factors that could have affected the VLBW of less educated, single black women during this period. The crack-cocaine epidemic is one possibility. In the largest population prevalence study ever undertaken, 29,494 women were tested for perinatal substances at 202 California hospitals in 1992. The percent of women exposed to cocaine at delivery was 13 times greater among black non-Hispanics (7.79 percent) than white, non-Hispanics (0.60 percent) and Hispanics (0.55 percent) (Vega et al. 1993). Importantly, urine assays used to test for cocaine have at most a 72-hour window. Thus, those

who are using cocaine within three days of delivery are likely to be highly addicted to not just cocaine, but tobacco, and may be trapped in a lifestyle that puts both mother and child at risk.

Brownell, M. D., M. J. Chartier, N. C. Nickel, D. Chateau, P. J. Martens, J. Sarkar, E. Burland, et al. "Unconditional Prenatal Income Supplement and Birth Outcomes." *Pediatrics* 137, no. 6 (June 1, 2016): e20152992–e20152992. <https://doi.org/10.1542/peds.2015-2992>.

The paper by Brownell et al. (2016) is a relevant citation because it looks at an unconditional income transfer for poor pregnant women in the province of Manitoba, Canada. The authors use inverse propensity score weighting (IPSW) to adjust their estimate. All women receiving Welfare in Manitoba are eligible to apply to the Healthy Baby Prenatal Benefit (HBPB), but not all do. Enrollment, therefore, is a choice and not exogenously assigned. Unless the authors can explain and control for why some women choose to enroll and others do not, the potential for selection bias is ever-present. HBPB provides \$80 CA per month in the second and third trimesters or approximately \$500 CA per pregnancy. The program lowered the risk of low birthweight by 29% and the risk of preterm birth by 24%. These are large effects that greatly exceed even the treatment on the treated estimates in HMS despite more than double the increase in income afforded by the federal EITC (HMS Online Appendix Table 3). Such large impacts recorded by HBPB may be related to selection bias given participation is a choice. For instance, low birthweight seems unaffected by confounders based on the sensitivity analysis conducted by the authors whereas babies that are small for gestational age (SGA) "...were very sensitive to unmeasured confounders.." (p. 5). We are not sure which confounders would affect SGA but have no effect on LBW, given the LBW includes SGA births. We appreciate the authors' straightforward presentation of the results, but the magnitude of the effects and the finding of possible confounding in a setting where patients chose to participate raises important concerns.