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: Cov(Y,Z)/Cov(X,Z) where Y is the outcome, Z the instrument and X the endogenous variable. They report a statistically significant reduced form {cov(Y,Z)] and a strong first stage {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 no 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.