# Poverty, the Earned Income Tax Credit, and the Mental Health of Children and Adolescents

Noah Kwicklis, 2017

**Abstract**

I hypothesize that cash transfers to poor households improve the mental health of recipient children. Specifically, I posit that the 1993 Omnibus Budget Reconciliation Act's expansion of the Earned Income Tax Credit (EITC) could have worked through a number of mechanisms to reduce the incidence of depression, anxiety, and antisocial behavior among children in eligible households, as reported by broad survey indexes. To test this claim, I estimate the intent-to-treat (ITT) effect of the EITC using a difference-in-differences (DID) identification strategy, with linear controls and household and region fixed effects. I find evidence that the federal tax credit expansion reduced externalizing[1] behavior and tendencies among low-income children.

*Keywords:* mental health, child behavior problems, externalizing behaviors, earned income tax credit

Research conducted for the completion of an undergraduate honors thesis at Cornell University. Many thanks to Hassan Enayati for advising the thesis, and to Francesca Molinari and Gregory Besharov for their guidance in its production. As always, any errors or oversights contained within are solely those of the author.

[1]Externalizing behavioral problems are those wherein the child "acts out" with aggressive or destructive behavior towards others and/or their surroundings (Liu 2004).

## 1. Introduction

According to the United States Census Bureau's 2016 report, 43.1 million Americans lived below the poverty line in 2015 – approximately 13.5% of the general population overall. Of this number, 14.5 million were children; the headcount poverty rate of the under-18 population that year stood at 19.7%. Additionally, studies conducted by the U.S. Centers for Disease Control and Prevention found that while nearly 8% of all Americans aged 12 and over reported signs of recent moderate-to-severe depression in 2014, the rate of reported depression symptomatology was over 15% among those living in poverty (and only 6% among those who were not) (Pratt and Brody 2014). The interaction of these two forces, childhood poverty and poor childhood mental health, may have particularly severe consequences for young people in low-income households with little access to treatment; Harvard University's Center on the Developing Child states that mental health problems in childhood often compound over time, causing yet greater issues as individuals grow older and enter adulthood.

Within this general context, a causal question may be formulated: Does poverty itself increase the incidence of poor mental health, specifically behavioral problems and affective mood disorders, among low-income children? And, if it does, can policymakers improve the welfare of the next generation while also simultaneously attaining other policy goals?

The primary causal question is difficult to definitively answer, as causal relationships between poverty and mental health and behavior are possible to intuit but challenging to empirically identify in a conclusive way. Income is endogenous to a range of variables that researchers cannot directly observe in most datasets, leading omitted variables to bias coefficient estimates away from true causal relationships. Furthermore, income and mental health are likely co-determined; an individual may become impoverished due in part to mental illness, or their mental illness may be exacerbated by factors linked to their impoverishment. Indeed, causality may run in both directions, such that mental illness causes poverty and poverty causes mental illness – resulting in a positive-feedback cycle. Attempts to estimate the extent to which one condition causes the other are thereby confounded by simultaneity, and untangling the second causal direction from the first is a nontrivial object of study.

Despite the complexity of the relationship between financial resources and mental well-being, a number of papers have investigated the first "mental health-to-poverty" direction of the interaction. Smith et al. (2010) suggest that individuals who grow up with psychological problems and affective disorders in childhood work fewer days out of the year and earn significantly less annual income in adulthood than their mentally well counterparts, after controlling for other health and socioeconomic factors. This finding by itself may be more correlative than causal; however, Mills et al. (2007) determine that programs designed to improve mental health conditions do appear to yield substantial positive returns to the labor force in the form of increased worker participation and productivity. Relatedly, Goetzel et al. (2004) estimate that depression and other mental illnesses cost employees and employers alike hundreds of dollars per afflicted individual per year (although the authors themselves are also cautious in the interpretation of their estimates). Furthermore, individuals with severe mental health problems may be unable to work at all, adding to the overall loss to the labor force. As such, if economic policy is at all tasked with improving the overall well-being and utility of society, then the imperative for policymakers to focus on problems of affective disorders like anxiety and depression is immediate and straightforward. Yet even in the purely pecuniary terms of dollars-and-cents, the potential aggregate drag of poor mental health on the national economy could be quite significant. Thus, for reasons pertaining to both citizen productivity and welfare, strategies to treat poor mental health should be of great interest to policymakers.

The converse causal direction, that of poverty being a driving factor in the incidence of poor mental health outcomes, is the main subject of this paper. It is also a line of inquiry where past policy may inform current research, and where current research may similarly inform future policy. The federal earned income tax credit (EITC) is one such case. As one of the largest anti-poverty programs in the country, the EITC is central to welfare policy in the United States. Like other tax credits, it is refundable; eligible households receive their full credit even if it exceeds their total tax liability, making it a negative income tax for many poor households. The Internal Revenue Service reports that approximately 26 million claimants received $65 billion in transfers through the EITC in the 2015 tax year, although roughly 20% of eligible taxpayers did *not* claim the benefit.

For those hoping to better understand the link between poverty and mental health, along with the effect of large government transfer programs upon the well-being of the public, systematic changes to the EITC represent substantial quasi-random "natural experiments," wherein changes to program benefits may deliver near-exogenous income shocks to select individual households. Namely, modifications to federal tax policies are often made in a manner independent of the choices of individual recipients, which helps break the link of individual action to benefit while preserving the link between benefit and individual action. This is especially useful from a research perspective when a policy is changed for one group but not another, even though the two groups are highly similar. Such was the case of the Omnibus Budget Reconciliation Act of 1993, hereafter referred to as OBRA93, which substantially increased the maximum EITC benefit available to low-income households with multiple children even as it left the benefits to similar single-child households largely unchanged. Using a difference-in-differences strategy, one may estimate the effect of this selective expansion to household financial resources on a variety of outcomes by comparing the changes experienced by the two groups before and after the policy.[1]

Yet even if a cash transfer "treatment" like the EITC is correlated with improved child mental health outcomes, the question of why increased financial resources might *cause* better mental well-being still remains. To this end, the EITC could act upon child mental health through a wide variety of channels. This paper is not itself concerned with proving and measuring each mechanism independently. Rather, it attempts to estimate the aggregate effect of all of them as the overall effect of the policy. Even so, a number of potential channels are reasonably plausible.

[1]Of course, households could still make changes to their behavior after a policy change is announced to opt into a higher benefit group. If the tendency to do so is correlated with unobserved traits that are in turn correlated with income and mental health, then endogeneity may persist. Even so, if the act of switching into a higher-benefit group is costly compared to the reward, the incentives of households to do so are arguably de-powered. The ensuing variation in household income is therefore, under certain assumptions and in the short term, subject to fewer endogeneity concerns on the individual level; the variation may then be used to identify a causal relationship between a change in financial resources and an otherwise endogenous outcome like mental health.

One possible mechanism by which wealth might influence mental health is through the latter's connection to nutrition. In O'Neil et al. (2014), a 12-study review of publications pertaining to diet and mental health, researchers found that children exposed to unhealthy diets (those high in saturated fats, sugars, and processed foods) consistently exhibited higher rates of anxiety and depression, although the authors were hesitant to assign a clear causal interpretation to the pattern, especially since the individual papers varied in quality and number of controls. Other studies examined specific nutrients in randomized control trial (RCT) experiments; in Nemets et al. (2006), a double-blind RCT found that children given higher levels of Omega-3 fatty acids, like those found in fish, exhibited lower levels of depression. Should expanded financial resources leads families to purchase higher-quality foods for their progeny, then these dietary changes could alter the biochemical conditions linked to affective disorders in youth. A cash transfer, through dietary means, could thus provide medical benefits. With this being said, it is also possible that unhealthy food is a normal good (as opposed to an inferior one) for lower-income households, as suggested in Kim and Leigh (2011), such that demand rises with income. If so, the magnitude of this particular wealth-health mechanism may be reduced for cash transfers up to a certain threshold.

Still, yet more mechanisms could be derived through improvements to access to better health care. Families with supplemented incomes may be better able to afford therapy, hospital bills, co-payments, and medications to a degree that would otherwise be unattainable. This greater access to mental and physical healthcare could boost child outcomes as well.

Moreover, problems caused by poor mental health affect not only the afflicted individuals themselves, but also whole clusters of their personal social networks. McLeod and Shanahan (1996) suggest that many affective mood disorders such as depression contain a socio-epidemiological component, whereby the negative mental health of parents can have an averse and lasting effect on the mental health of offspring (although the causality may again be complicated). If the EITC improves the mental health of parents, as it has been shown to do so for mothers in Boyd-Swan et al. (2013) and Evans and Garthwaite (2014), then parental benefits may be passed on into child benefits as well, causing improvements in parent outcomes to improve child outcomes. In as many words, children whose parents are in better mental health could themselves face fewer adverse pressures.

Underlying all of these channels is the fact that children in poor households are almost always subject to the same financial and resource constraints as their parents; these pressures by themselves may be chronic stressors with harmful long-term effects for all parties involved. As previously alluded to, children are over-represented in poverty; while minors represent less than a quarter of the national population, they constitute over a third of all impoverished individuals (Proctor et al. 2016). They thus run the risk of being disproportionately exposed to poverty-driven mental health problems as well. Additionally, for a young person, even small initial alterations in their rate of human and physical capital accumulation can compound over the course of their lifetime, radically altering their overall well-being and the trajectory of what Cunha et al. (2005) term their "life-cycle skill formation." Therefore, when considering the immediate and long-term welfare implications of antipoverty programs, children and adolescents should be considered a key demographic, as even small investments in transfer payments today could result in sizable changes to the aggregate welfare of society in the future.

As such, for this paper, I study the interaction between poverty and the mental health of minors, to which the following two questions are central:

1. How did the Omnibus Budget Reconciliation Act of 1993's expansion of the earned income tax credit influence the onset of mental health and behavioral problems amongst youth living in low-income households when other factors such as age, sex, race, parental marital status, caregiver labor decisions, and unobserved (but fixed) effects are controlled for? Did the substantial increase in the transfer payments allotted to multi-child families lead to greater improvements in mental health outcomes when compared to single-child households, whose EITC benefit was raised by a smaller amount?

2. How do household resources affect the mental health and behavior of poor children, if at all?

In doing so, I hope to estimate the efficacy of supplemental income programs as a supporting treatment for poor mental health amongst some of the most disadvantaged members of society. If effective in this way, such programs like the EITC could improve mental health in children by improving their (and their parents') access to food, housing, and healthcare, thereby reducing environmental factors conducive to chronic mental health problems. I hope to add to the literature concerned with the non-financial outcomes of economic antipoverty policy and the earned income tax credit. But first, a discussion of this existing literature is warranted.

## 2. Household Behavior and the EITC

There is already a significant and growing body of research committed to investigating the effects of the EITC on a wide variety of outcomes. Much of it is directed toward studying the labor-supply decisions of its recipients; many papers have found that unmarried mothers tend to respond to the tax credit by increasing their labor supply, although the response among married women appears to be more muted (Evans and Garthwaite 2011, Eissa and Hoynes 2006). As a general test of my research design, I replicate these findings in Appendix A of this report. However, many more recent papers are now seeking to analyze the *non-financial* effects and externalities of the earned income tax credit, including those which pertain to child development and claimant health. This paper intends to contribute to this common literature by further exploring the connection between financial and mental well-being.

Among the EITC studies pertaining to psychological and general health, Evans and Garthwaite (2011) and Boyd-Swan et al. (2013) both appeared to find improved health and mental health outcomes among adult women with children after the expansion of the EITC. If maternal mental health influences child health at all, then one might expect to find that the program yields positive influences on child mental health as well. Moreover, greater and more stable access to food, shelter, and health care could also positively impact children. These possibilities appear to be generally confirmed in studies by Averett and Wang (2015) and Hamad and Rehkopf (2016), who both found that broad aggregate measures of child non-cognitive development and home environment were improved by the transfer payment programs, although they did not delve into specific sub-aspects of child mental health, as this paper does. Milligan and Stabile (2008), meanwhile, looked at specific indexes of aggression and anxiety amongst Canadian children and found that they were improved following child tax credit expansions; female children in particular appeared to receive greater mental health benefits from the Canadian program.

Other specific benefits for children from the EITC are detailed in Dahl and Lochner (2012) and Maxfield (2013); although their papers did not explicitly focus on mental health and happiness, they did find that child school performance was improved by the program. If measures of child cognitive achievement are improved by the EITC, it is possible that the policy may improve emotional and psychological ones as well.

With this being said, one might still question if the EITC actually provides a "treatment" by which households may be affected at all, as different welfare programs can interact with one another to affect eligibility. Maxfield (2014) used a family-size difference-in-differences model to show that the EITC does significantly improve the net financial standing of households, and that it also encourages maternal participation in the labor force. Although these findings confirm that the EITC improves family financial resources, they could also suggest that the effect of the program on child mental health is ambiguous, as parents could choose to work increasingly long hours, potentially complicating the emotional health of their progeny even as the family's resources improve. Indeed, economic theory would indicate that the EITC provides households with a strong incentive to enter the labor force, although this incentive may be reduced for higher earners whose benefits are phased out with additional income (Hotz and Sholz 2003). The complexity of agent behavior and reaction to policy may thus confound straightforward interpretations of the EITC as a simple positive income shock, unless the labor participation effects of the policy are accounted and controlled for.

Of these studies, Averett and Wang (2015) and Hamad and Rehkopf (2016) are the most similar to this paper. Both studied the EITC's OBRA 1993 expansion and used data from the National Longitudinal Survey of Youth 1979 Children and Young Adult supplement. In particular, they studied the effects of the EITC on the aggregate behavioral problems index (BPI) and the home environment index. Both papers therefore examined the same research question using the same dataset, albeit with slightly different methodologies. Neither paper, however, delved more deeply into the individual questions and sub-indexes contained within the BPI, a task which I plan to undertake. Milligan and Stabile (2008) did delve into these specific questions, but they used a slightly different policy (the Canadian child tax credit) in a different nation using a different dataset using an instrumental variable identification strategy. Therefore, this thesis aims to replicate and expand upon the respective findings of these papers by delving into the more detailed changes in reports of anxiety, antisocial behavior, and depression in the United States in response to the earned income tax credit.

Within this EITC literature, it is important to note that the actual transfer payments to households from the EITC are not the independent variable onto which key outcomes are regressed. Rather, each study measures only what Boyd-Swan et al. (2013) call the "intent-to-treat," or "ITT"; the actual EITC payments received by the household are unobserved, but the *eligibility* of the household to receive payments from the tax credit may be reconstructed from survey questions. As Boyd-Swan et al. (2013) note, estimates of the ITT impact avoid biases like the additional self-selection into treatment that could occur when families choose to claim or not claim the credit for which they are eligible. However, as I mentioned in the introduction of this paper, the IRS reports that only some 80% of EITC-eligible taxpayers claim their full benefit. Thus, attempts to interpret the ITT estimate as the effect of actually receiving the transfer are likely to inadvertently bias the estimate toward zero (Hamad and Rehkopf 2016). Even so, the estimate may still accurately describe the real policy impact of expanding the credit, if 20% of eligible individuals consistently expected do not to opt into the program.

## 3. Data

### *3.1. The NLSY and the Behavior Problems Index (BPI)*

To conduct my analysis, I use the National Longitudinal Survey of Youth 1979 (NSLY79) and the NSLY79 Children and Young Adults (NLSCY) supplementary dataset; the latter survey was first administered to the progeny of the participants of the original NSLY79 in 1986. These datasets are both widely used in EITC studies, as they provide detailed biennial snapshots of the socioeconomic status and well-being of a given respondent and their household over the course of several years. The data sets are also complementary; children from the NLSY Child can be matched with their mothers from the NLSY79, allowing researchers to gain a broad profile of both the child and their home environment. Most of the variables of each dataset are available for public use, although some sections (like those that pertain to the geographic location of respondents) are confidentially maintained by the Bureau of Labor Statistics to ensure that individual respondent identities cannot be compromised. This paper uses data from both only the public access section.

Within my study, I focus on the Behavior Problems Index (BPI) and its six sub-indexes: the anxiety and depression, immature dependency, headstrong/oppositional defiance, hyperactivity, antisocial behavior, and peer conflict scales. These scores are designed to provide a general insight into the non-cognitive functioning of each child aged 4 and over, and are calculated using the trichotomous answers to 28 questions provided by the mother. To each individual question about their child, the parent can answer "Never True," "Sometimes True," and "Often True;" an answer in the latter two adds an additional point to the BPI overall index and the relevant subindex. These indexes are thus *dichotomously* generated, as answers to each of their component questions are reduced to either a zero or one and then tallied.

In addition to the six dichotomously-generated sub-indexes, the NLSY maintains a trichotomously generated BPI index and two trichotomously generated sub-indexes as well. The first sub-index roughly gauges the respondent child's proclivity to externalize their behavior; in the field of child psychology, externalized behavior often manifests as increased aggression, hyperactivity, and/or antisocial tendencies, where the child "acts out" on those around them (Liu 2004). The second sub-index generally measures the child's tendency to internalize behaviors– that is, become depressed, withdrawn, and/or anxious. These two trichotomous sub-indexes thereby measure behavioral problems with greater specificity than the overall BPI, but with less specificity than the six dichotomous sub-scales. Since these indexes are trichotomous, a "Never True" answer is coded as a 0, a "Sometimes True" answer is coded as a 1, and an "Often True" answer is coded as a 2, such that the tally constitutes the overall index score.

Of course, a parent can also refuse to answer a question about their child altogether. My study is therefore predicated on the proposition that parents whose children are 14 years of age or younger are willing and able to adequately assess and report the mental health of their progeny to the NLSY, or at least mis-report in a way that does not entirely obfuscate a response from their child to improved household resources. It is possible that stigmas related to mental illness could call this assumption of adequate reporting into question. Still, since the NLSY goes to great lengths to ensure respondent anonymity, parents likely face fewer social pressures in this regard. I also assume that parental reporting of behavioral problems is not significantly affected by cash transfers. If families interpret the behavior of their children more charitably following a transfer payment, for instance, then they might also report a reduction in behavior problems when no such reduction actually occurred. This could then be misinterpreted as an improvement in child behavior caused by the policy, when all that was actually changed was the surrounding adults' perception.

For studies dealing with federal policy roll-outs, the timing of each wave of the NLSY can also prove quite consequential. The fielding periods for the NLSY Child surveys typically last between six to eight months during the survey year, wherein respondents are asked questions pertaining to their activities from the previous calender year (NLSY). For instance, households in 1994 were surveyed from June to December. The initial EITC expansion of OBRA93 was also implemented that year, and while households typically file their tax returns in April, the returns can take up to three weeks to be processed (IRS 2016). Additionally, the IRS reports that it is legally mandated to facilitate EITC refunds until the following February. As such, some of the households surveyed by the NLSY in 1994 could have already claimed and received their tax credit, while others also surveyed that year could have claimed their benefit following their interview. Due to this ambiguity, I have dropped 1994 from the years under explicit consideration in my dataset.

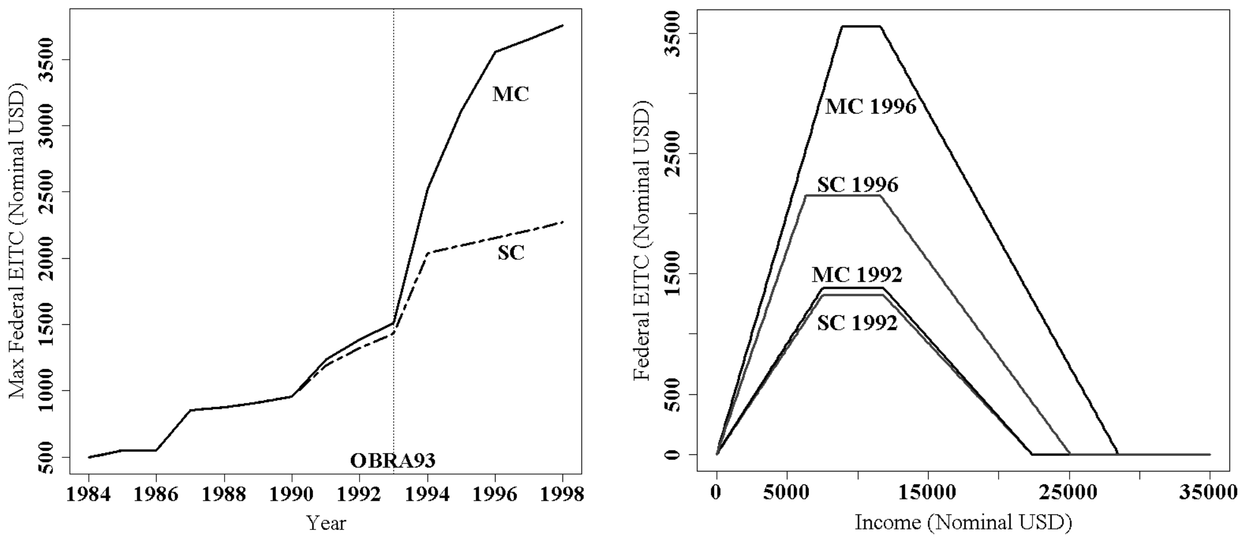
Lastly, the NLSY79 itself notes that it oversampled low-income households and households of color.[2] Thus, the external validity of findings generated from the sample could in theory be questionable, should the findings be extrapolated to the entire United States population, although the NLSY provides a range of observation weightings to correct for this. Still, most anti-poverty transfer programs are geared toward low-income families, so any results from my project should not be significantly out-of-sample when applied to policy evaluation. It should also be noted that when it comes to interpreting the EITC as a broad-base exogenous income supplement, families with parents too disabled or mentally ill to work (despite whatever incentives offered to them) will not be eligible for the EITC and cannot opt to become eligible, as the benefit is contingent upon hours worked. These households had no variation in EITC benefits capable of being used as identifying variation in household income. The external validity of my results when applied to this particular group of households is therefore uncertain as well.

### *3.2. OBRA93 and Group Assignment Within the Sample*

Since the NLSY79 sample was primarily composed of individuals from lower-income groups, many of the survey's respondents were likely affected by the Omnibus Budget Reconciliation Act of 1993 (OBRA93). The Act dramatically increased the federal earned income tax credit over the course of 1994 and 1995, in what would become the largest expansion

[2]Interestingly, the main racial categories included in the NLSY79 Child supplement were limited to "black," "Hispanic," and "non-black, non-Hispanic" (although more detailed information pertaining to ethnicity was gathered by the main survey). As such, more nuanced racial identities and social dynamics may be overlooked by analyses that partition the data based on these three categories alone.

**Figure 1:** The Earned Income Tax Credit and OBRA93




**(a)** EITC Maximum Benefit by Year **(b)** EITC Benefit by HH Income, Type and Year

Figures generated by the author using the NBER's TAXSIM package. **SC** denotes simulated benefits to single-child households, while **MC** denotes simulated benefits to multi-child households.

in the history of the program (Averett and Wang 2015). This was by no means the first expansion of the EITC; the credit has been available to low-income wage-earning households with children since the late-1980s, and has grown steadily in size since its inception (Evans and Garthwaite 2011). This is illustrated in Figure (1.a), which was generated using values simulated by the TAXSIM program, a resource for simulating historical tax obligations maintained by the National Bureau of Economic Research. For many households, OBRA93 induced only a relatively modest change to their annual financial resources; a single-child household making \$10,000 a year saw their maximum benefit increase from \$1324 in 1993 to \$2,152 in 1996, for a total change of about \$828 in the maximum benefit.

Households with multiple children under the age of 19, however, benefited far more substantially. Similar to their single-child counterparts, the maximum credit for multi-child households in 1992 was \$1384 according to TAXSIM estimates. By 1996, however, their maximum benefit had risen to over \$3,556, for a difference of \$2,272; multi-child families thus received an expansion from OBRA93 that was nearly double that of single-child ones. The total difference-in-differences in the maximum benefit, the degree to which the maximum earned income tax credit was differentially expanded for multi-child households in excess of single-child ones, was \$1,444.

These changes to the EITC for both household types, along with the EITC's benefit conditional on income, are depicted in Figure (1.b). The \$1,444 differential in maximum benefit that opened between single-child and multi-child households after OBRA93 may be construed as a kind of "treatment" delivered by the policy to the larger household type, but not to the smaller one. It is therefore useful to introduce a quasi-experimental language to the discussion of the policy analysis: if the over-\$1,000 differential increase in transfers to multi-child families is the "treatment" under consideration, then multi-child households may be interpreted as a kind of treatment group, while single-child low-income households serve as a kind of control group. Using a difference-in-differences methodology, the changes over time of the two groups may thus be compared to estimate the effect of a cash transfer of up to $1,444 on recipient outcomes, provided certain identifying assumptions. However, one may note that OBRA93 did *not* generate a uniform treatment for all members of the treatment multi-child group; households making more or less than the benefit-plateau incomes depicted in Figure (1.b) likely received a reduced treatment compared to those in the maximum-benefit plateau range. Since some households on the margins of the eligible income group only received very small transfers, my estimates of the effect of these transfers could be somewhat biased toward zero.

The precise demarcation of the EITC-eligible group under consideration is also potentially fraught. As a large number of papers have shown, the expansion of the conditional tax credit likely influenced household labor supply decisions. Therefore, if households are assigned to the EITC-group based on their level of income or credit eligibility in the year of interview alone, a problem of endogenous selection could emerge. For instance, the extra benefits of the EITC may encourage householders to work more hours for augmented pay. A household in the EITC-eligible group in 1992 could thus have responded to the OBRA93 expansion by working more hours and gaining greater work experience and income. They could then have removed themselves from EITC eligibility by the date of their interview with the NLSY in 1996 – in response to the tax credit itself. To avoid the problem of households removing themselves from the treatment group in response to treatment, other EITC papers including Evans and Garthwaite (2011) and Averett and Wang (2015) sort households into the EITC group based on the education level of the female householder at the time of interview, under the assumption that level of education is not as responsive to the tax credit as income or predicted credit eligibility. This is usually done by sorting households where the mother has a high school diploma or less into the likely EITC-eligible group, while households with college graduate mothers are placed into the likely EITC-ineligible group.

Rather than cutting the data by maternal educational attainment, I use an alternative (non-dynamic) measure of socioeconomic status to assign families into the likely-eligible and likely EITC ineligible groups: the household's total income in 1992, on the eve of the OBRA93 expansion. More precisely, I place households who made less than or equal $30,000 into my likely EITC-eligible sample, and those households that made more in 1992 into my likely ineligible sample – this latter group is then omitted from my study[3]. This static criterion does not allow households to flux in and out of the sample as a response to the policy.

I therefore group my NLSY Child observations into 2 categories: likely EITC-eligible multi-child households (the treatment group) and likely EITC-eligible single-child households (the control group). Since each of these groups has its own before-OBRA93 period and after-OBRA93 period, my control and treatment groups split the sample into 4 groups in total. The TAXSIM-estimated EITC-eligibility statistics for each are displayed in Table 1, along

[3]Theoretically, these wealthier households could be used as an additional set of control groups, for a triple-differences estimator. However, the NLSY Child

**Table 1:** Average EITC Benefits to Children by Household Income in 1992 and Number of Children, Before and After OBRA93

| | Single-Child | | Multi-Child | |
|---|---|---|---|---|
| | Before | After | Before | After |
| *Household Income in 1992* $\leq \$30,000$ | | | | |
| Average EITC Benefit | 74.25 | 99.14 | 136.30 | 371.25 |
| Eligible for the EITC (%) | 12.3 | 10.4 | 20.7 | 21.0 |
| Average EITC Benefit Among Recipients | 603.09 | 955.71 | 659.53 | 1,771.06 |
| Observations | 601 | 482 | 4,166 | 2,953 |

Estimated using the NBER TAXSIM package.

for the total number of observations in each category.

Given this grouping of sample observations, it is clear that multi-child households in the likely-EITC eligible households received a substantially larger increase in benefits than single-child eligible households as simulated by the TAXSIM program, both on average across the entire respective group and amongst the households actually predicted to have received a benefit. Still, many households in this "likely-eligible" group likely received no earned income tax credit at all, as they did not report any wage income in the given year.

Another (potentially more serious) source of endogenous self-selection, however, could be driven by households changing their number of dependent children, which would cause them to switch between the control and treatment groups. As alluded to in the introduction, the statutory implementation of the OBRA93 expansion spread was not induced by individual household choices. However, it is possible that childless or EITC-eligible single-child households could have opted into the higher expansion after the fact, albeit at the relatively high cost of having an additional child. Their reasons for doing so might be unobserved, but they will still be correlated with treatment group assignment, as treatment group assignment is predicated on household size. If these variables are 1) omitted and relegated to the error terms of the econometric models discussed in the next section of this paper, 2) correlated with treatment group assignment (by construction), and 3) also correlated with child outcomes through channels other than the cash transfer, then the coefficient estimates calculated using the EITC expansion may no longer identify the true causal effect of the policy on the recipient children's well-being. In fact, (Baughman et al. 2003) does find that expansions to the EITC may have inadvertently encouraged previously childless individuals to have children – to an economically small but still statistically significant degree. Even so, the authors stress that the effect is minute for most practical purposes, and a following paper (Baughman and Dickert-Conlin 2009) suggests that the EITC did not have a clear pro-fertility effect at all.

On the following page, I provide summary statistics describing children from each of the 4 groups within my NLSY Child sample. Each group is almost evenly split between male and female children. The average child from the likely EITC eligible group is about about 9 years old in 1992, and the sample did not predictably age with time, since children younger than 4 and adolescents older than 14 were constantly aging into and out of the NLSY Child sample with each biennial survey round. As one might expect, NLSY children from likely EITC-eligible households also received far higher foodstamp, Aid to Families with Dependent Children (AFDC), and social security insurance benefits than children from higher-income households, as reported by their parents. Given the potential differential variation of these programs across groups with time, especially in the case of changes to AFDC benefit, I attempt to control for the effect of these fluctuations upon outcomes by including the dollar-value of these transfers as linear terms in my regressions.

**Table 2:** Likely EITC-Eligible (1992 Household Income $\leq$ \$30,000) Child
Sample Means and Percentiles by Year

| | Panel A: Single-Child | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Before OBRA93 | | | | After OBRA93 | | | |
| | Mean | P25 | P50 | P75 | Mean | P25 | P50 | P75 |
| Child Age | 9.09 | 6.00 | 9.00 | 11.0 | 9.92 | 7.00 | 10.0 | 13.0 |
| Identified Male | 0.483 | 0 | 0 | 1.00 | 0.488 | 0 | 0 | 1.00 |
| Identified Black | 0.376 | 0 | 0 | 1.00 | 0.353 | 0 | 0 | 1.00 |
| Identified Hispanic | 0.206 | 0 | 0 | 0 | 0.180 | 0 | 0 | 0 |
| Parent Age | 30.3 | 28.0 | 30.0 | 32.0 | 35.8 | 34.0 | 36.0 | 38.0 |
| Mother Married | 0.268 | 0 | 0 | 1.00 | 0.367 | 0 | 0 | 1.00 |
| HH Wage Income | 9489 | 0 | 9600 | 16000 | 15029 | 5000 | 14000 | 22000 |
| No Children in HH | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Mother Employed | 0.654 | 0 | 1.00 | 1.00 | 0.795 | 1.00 | 1.00 | 1.00 |
| Mother's Wkly Hrs Wrked | 26.1 | 7.85 | 31.7 | 40.0 | 31.2 | 19.5 | 39.5 | 40.0 |
| Mother's Hrly Pay | 5.74 | 0 | 5.29 | 8.27 | 8.80 | 4.23 | 7.24 | 10.7 |
| AFDC Benefit | 685 | 0 | 0 | 0 | 291 | 0 | 0 | 0 |
| Foodstamp Benefit | 375 | 0 | 0 | 0 | 212 | 0 | 0 | 0 |
| SSI Benefit | 177 | 0 | 0 | 0 | 246 | 0 | 0 | 0 |
| Observations | 601 | | | | 482 | | | |

| | Panel B: Multi-Child | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Before OBRA93 | | | | After OBRA93 | | | |
| | Mean | P25 | P50 | P75 | Mean | P25 | P50 | P75 |
| Child Age | 9.32 | 7.00 | 9.00 | 12.0 | 9.76 | 7.00 | 10.0 | 12.0 |
| Identified Male | 0.500 | 0 | 0.500 | 1.00 | 0.511 | 0 | 1.00 | 1.00 |
| Identified Black | 0.440 | 0 | 0 | 1.00 | 0.385 | 0 | 0 | 1.00 |
| Identified Hispanic | 0.247 | 0 | 0 | 0 | 0.230 | 0 | 0 | 0 |
| Parent Age | 30.5 | 29.0 | 31.0 | 32.0 | 35.4 | 34.0 | 35.0 | 37.0 |
| Mother Married | 0.416 | 0 | 0 | 1.00 | 0.531 | 0 | 1.00 | 1.00 |
| HH Wage Income | 5447 | 0 | 1200 | 10000 | 9910 | 0 | 7000 | 16500 |
| No Children in HH | 3.05 | 2.00 | 3.00 | 4.00 | 2.99 | 2.00 | 3.00 | 4.00 |
| Mother Employed | 0.493 | 0 | 0 | 1.00 | 0.637 | 0 | 1.00 | 1.00 |
| Mother's Wkly Hrs Wrked | 17.3 | 0 | 11.9 | 35.0 | 22.9 | 0 | 25.0 | 40.0 |
| Mother's Hrly Pay | 3.80 | 0 | 2.00 | 6.13 | 6.25 | 0 | 5.29 | 9.13 |
| AFDC Benefit | 1558 | 0 | 0 | 2760 | 755 | 0 | 0 | 0 |
| Foodstamp Benefit | 1336 | 0 | 0 | 2760 | 777 | 0 | 0 | 916 |
| SSI Benefit | 453 | 0 | 0 | 0 | 430 | 0 | 0 | 0 |
| Observations | 4,166 | | | | 2,953 | | | |

Estimated using data from the NLSY79 Child Supplement P25 denotes the value of the variable at the lower 25th percentile of the sample distribution, P50 describes the variable's median, and P75 denotes the variable's upper 25th percentile Wages, pay, and benefits denominated in current U.S dollars.

## 4. Econometric Models

To adopt the quasi-experimental language of (Rubin 1980), let $Y_{it}$ denote the mental health or behavioral outcome of child $i$ at time $t$. From there, I define $D_{it}$ as an indicator that takes a value of 1 if child $i$ receives a treatment in period $t$, and zero otherwise. Since the multi-child maximum EITC benefit was expanded by OBRA93 at a greater rate than the single-child one, the "treatment" here is the EITC expansion for which multi-child families were eligible that single-child families were not. In the interest of uncovering the average causal effect of the intervention, I now define $Y_{it}^1$ as the outcome of the child with treatment, while $Y_{it}^0$ is the outcome of the child in the intervention's absence.

By construction, the true expected intent-to-treat effect of the EITC on a mental health outcome of child $i$ in time $t$ is simply the expected outcome of the child had they received the treatment, minus the expected outcome of the child had they not.

$$\delta_{True} = \mathbb{E}\left[Y_{it}^1 - Y_{it}^0\right] \tag{1}$$

However, one of the terms in (1) for child $i$ will be counterfactual by necessity, as each child either did receive or did not receive treatment in time period $t$ – but not both simultaneously. For instance, for a child treated at time $t$, the first term $Y_{it}^1$ will be observed in the data, while the term on the $Y_{it}^0$ will not. In a more straightforward experiment where treatment is as-good-as-randomly assigned, a researcher might use the expected outcomes of children from non-treatment (single child) families to approximate the counterfactual outcomes of children in the treatment group in a treatment period (multi-child households after 1993). Alternatively, one might also use the outcomes from treated children in a time period before the treatment took place, to the same purpose. Either of these could then be substituted into the second term of (1) to find a first-difference. However, multi-child families may inherently have different traits than single-child families, with different baseline levels of mental health. Additionally, trends in outcomes with time could also make treated child outcomes in prior years unsuitable stand-ins for the counterfactual term.

To account for these imperfections in randomization and possible secular time trends, a difference-in-differences (DID) strategy may be applied; this strategy effectively applies both of the aforementioned first-difference methods concurrently, as the name would suggest. Let $MC_i = 1$ indicate that a child is in a multi-child household, and let $MC_i = 0$ denote that the child is in a single-child household. The effect of the policy on children from multi-child households may thus be obtained by first calculating the following two differences, where the subscript $Pre$ indicates a pre-OBRA93 observation and $Post$ denotes a post-OBRA93 observation:

$$\delta_0 = \mathbb{E}\left[Y_{i,Post}^0 | MC_i = 0\right] - \mathbb{E}\left[Y_{i,Pre}^0 | MC_i = 0\right]$$

$$\delta_1 = \mathbb{E}\left[Y_{i,Post}^1 | MC_i = 1\right] - \mathbb{E}\left[Y_{i,Pre}^0 | MC_i = 1\right]$$

These may then be used to calculate the difference-in-differences estimate:

$$\delta_{DD} = \delta_1 - \delta_0$$

All of the terms in $\delta_0$, $\delta_1$, and $\delta_{DD}$ may be directly estimated from the data; none of them are counterfactual. Suppose it holds that the average change in single-child outcomes is equal to what the average change in multi-child outcomes *would have been if not for the policy*, such that the average no-treatment delta is the same regardless of group assignment:

$$\mathbb{E}\left[Y^0_{i,Post} - Y^0_{i,Pre}|MC_{it}\right] = \mathbb{E}\left[Y^0_{i,Post} - Y^0_{i,Pre}\right]$$

From this mean independence condition,

$$\delta_0 = \mathbb{E}\left[Y^0_{i,Post} - Y^0_{i,Pre}|MC_i = 0\right] = \mathbb{E}\left[Y^0_{i,Post} - Y^0_{i,Pre}|MC_i = 1\right]$$

The DID estimate $\delta_{DD}$ then collapses to both the ITT average effect-on-the-treated and the ITT average effect:

$$\begin{aligned}\delta_{DD} =&(\mathbb{E}[Y^1_{i,Post}|MC_i = 1] - \mathbb{E}[Y^0_{i,Pre}|MC_i = 1])\\ &- (\mathbb{E}[Y^0_{i,Post}|MC_i = 1] - \mathbb{E}[Y^0_{i,Pre}|MC_i = 1])\\ =&\mathbb{E}[Y^1_{i,Post}|MC_i = 1] - \mathbb{E}[Y^0_{i,Post}|MC_i = 1] = \mathbb{E}\left[Y^1_{i,Post} - Y^0_{i,Post}\right]\end{aligned}$$

Which is simply a reformulation of (1). As such, the difference-in-differences coefficient $\delta_{DD}$ will be an estimate of the true intent-to-treat effect of an EITC expansion, under the identifying assumption that the counterfactual trend of the multi-child treatment group is well-modeled by the trend of the single-child control group. This $\delta_{DD}$ coefficient may be equivalently estimated using the unadjusted DID regression

$$Y_{it} = \beta_0 + \beta_1 MC_i + \beta_2 After_t + \delta_{DD}(MC_i \times After_t) + \varepsilon_{it} \tag{2}$$

Where $After_t = 1$ if the observation is taken from a year after the EITC expansion and zero otherwise (equivalently, the dummy variable is the indicator function $After_t = \mathbb{1}\{t > 1993\}$). Similarly, $MC_i = 1$ if child $i$ belongs to a multi-child household and 0 when they belong to a single-child household. To keep treatment effect coefficients comparable across models, I standardize $Y_{it}$ and regress onto the z-score of the dependent variable during estimation.

But is the identifying assumption, the parallel trends assumption, truly safe to make? It cannot be tested directly in a treatment year, as the treatment itself is measured by the deviations from parallel trends that it induces. Still, one may observe the trends in treatment and control outcomes in time periods prior to the treatment to see if the parallel trends assumption was valid in at least pre-treatment years. This is done more rigorously in the Results section of this paper. As a first-pass analysis, however, I plot the average mental health outcomes of children from single-child and multi-child households by year. These graphs are presented in Figure 1*b*. For many of the affective and behavioral indicators, the identifying assumption appears relatively sound. The average overall behavior problems scores trended together fairly consistently until 1993, after which point the two diverged. The assumption appears similarly sound for the internalizing behavior, immature emotional dependence, and antisocial behavior sub-indexes. For other outcomes, the parallel trends assumption may be somewhat arguable.

Using NLSY data from 1990 to 1998, the baseline unadjusted DID regression may be further modified to contain a variety of controls:

$$Y_{ijst} = \beta_1 MC_i + \beta_2 After_t + \delta_{DD}(MC_i \times After_t) + X'_{ijt}\theta + \eta_j + \mu_s + \lambda_t + \varepsilon_{ijst} \tag{3}$$

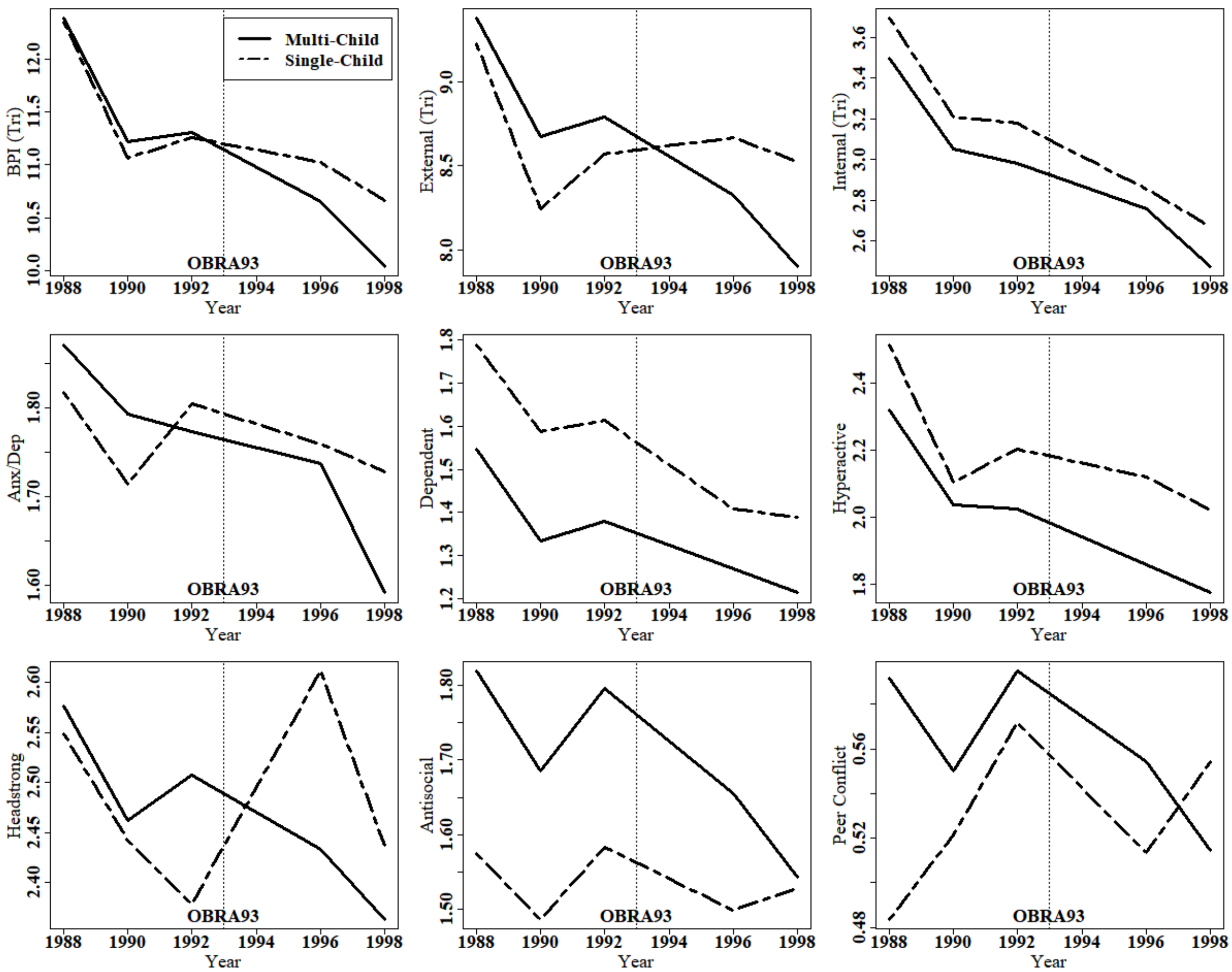



**Figure 2:** Outcome Trends Among Likely-EITC Eligible Children, Before and After OBRA93

Figures generated by the author using data from the NLSY Child supplement. Dashed lines denote the average outcomes of children from single-child households, while solid lines denote the average outcomes of children from multi-child households.

Specification (3) now includes $X_{ijt}$, a vector of maternal, household-level, and child-level controls. Child controls consist of dummies for the child's age and gender, along with interacted dummies for the child's age and year of interview. Parental controls include dummies for the mother's age and marital status, along with linear terms for the household's total yearly wage income, the number of weekly hours the mother works, the mother's hourly pay, and alternative forms of welfare received by the household that year including AFDC, foodstamp, and social security insurance benefits. Additionally, to reduce household-level omitted variable bias, $\eta_j$ captures household fixed effects (associated with mother $j$); any time-invariant factors not explicitly included in the model (like race, family history, and persistent household dynamics) are absorbed by this term, which is equivalent to a dummy variable for each of the households in the sample. Additionally, $\mu_s$ denotes a region-specific fixed effect for factors related to mental health and treatment program exposure that are homogeneous within broad geographies but heterogeneous across them. Finally, I include an additional year fixed-effect $\lambda_t$ to account for general macroeconomic trends affecting all children in each year within the pre- and post periods.

With these additional covariates, the identifying assumption necessary is then one of *controlled* parallel trends, conditioned on observable variables:

$$\mathbb{E}[Y^0_{i,Post} - Y^0_{i,Pre}|\ MC_i,\ X_{ijt},\ \mu_s,\ \lambda_t,\ \eta_j] = \mathbb{E}[Y^0_{i,Post} - Y^0_{i,Pre}|\ X_{ijt},\ \mu_s,\ \lambda_t,\ \eta_j] \quad (3^*)$$

In other words, after adjusting for family, state, and year fixed effects, and after linearly controlling for child and household characteristics, the average change in child outcomes of the treatment (multi-child) and control (single-child) groups must have been equal in the absence of of intervention for (3) to identify the true effect (1) conditioned on observables – assuming the linear specification of (3) is correct.

Although this model takes data from 4 biennial waves of the NLSY child survey, it pools the first two rounds (1990 and 1992) into a pre-OBRA expansion and the latter two (1996 and 1998) into one large post period. By these means, it estimates a single measure of the expansion's average effect in the years after the expansion has taken place.

However, a regression model like (3) assumes that the residuals of all of the observations are independently and identically distributed, such that the variance-covariance matrix of the model's residuals is diagonal and scalar: $\text{Cov}(\varepsilon) = \sigma^2_\varepsilon E$, where $E$ is the identity matrix. This may not be true, however, if a child's mental health in one year is serially correlated with their mental health in another year. Additionally, it is also likely that the mental health and behavior of siblings are correlated as well, due to shared family history, home environment, and the strength of sibling interactions; these observations may be correlated across periods as well. Regression (3) does contain a mother fixed effect term, effectively removing the mean from household clusters, but, as Angrist and Pischke (2009) note, this within-cluster transformation may not remove autocorrelation within the cluster itself. To reduce the risk of Type I error, I therefore cluster the model's robust standard errors on the family level to allow for arbitrary correlations within observations from the same family, a strategy that should asymptotically correct for both general forms of heteroskedasticity and within-family intertemporal autocorrelation. Since my number of family clusters and my number of observations are both quite large, the asymptotic assumptions of this estimator are reasonably justified.

All of the models discussed so far assume a single treatment effect of the OBRA93 expansion on the multi-child group in the post period. However, this may not necessarily be the case. After financial resources of a household are increased, it could theoretically take some time, even multiple years, for the effect of the material change to substantially alter the outcomes of recipient individuals. This is particularly the case where the effect of the treatment compounds with each transfer. To account for this possibility in the difference-in-differences model, I can interact the year-specific dummies of the time fixed effect with the multi-child binary variable to yield a multilevel regression with one treatment group, one control group, but four time periods:

$$Y_{ijst} = \beta_1 MC_i + \sum_{\tau=1992}^{1998} \lambda_\tau \mathbb{T}(\tau) + \sum_{\tau=1992}^{1998} \delta^\tau_{DD}(MC_i \times \mathbb{T}(\tau)) + CX'_{it}\gamma + HX'_{it}\theta + \eta_j + \mu_s + \varepsilon_{ijst} \quad (4)$$

Where $\mathbb{T}(\tau) = \mathbb{1}\{\tau = t\}$, an indicator function that generates dummy variables for each year (note that 1990 is the base year, for which all $\mathbb{T}(\tau) = 0$). As such, the regression models an outcome for child $i$ given mother $j$ in state $s$ at time $t$; we may additionally condition the observation on $c$ to indicate the child's status in a multi-child household type to replace the dummy $MC_i$ with a "fixed effect" for household type, since the two-level fixed effect and a dummy variable are equivalent (although the former concept is more easily generalized than the latter). By these means, (4) may be written more compactly as

$$Y_{ijcst} = \alpha_c + \lambda_t + \delta_{DD}^t(MC_i \times \mathbb{T}(t)) + CX_{it}'\gamma + HX_{it}'\theta + \eta_j + \mu_s + \varepsilon_{ijcst} \tag{4'}$$

The crucial distinction between (4) and (3) is that instead of estimating one treatment effect for all of the post-OBRA93 years, regression (4) calculates a difference-in-differences estimate of the effect of the OBRA93 expansion in *each* year, denoted $\delta_{DD}^t$, with "leading" and "lagged" effects (Angrist and Piscke 2009). While the identifying assumptions remain the same (see 3*), this more flexible model makes it possible to observe how the estimated effect of the policy changes with time (Pischke 2005).

As previously stated, model (4) is advantageous over (3) in its capacity to distinguish between different specific years, albeit at the cost of reduced degrees of freedom. However, it may additionally assist in a falsification test of the original model: in the absence of anticipatory effects of the 1993 EITC expansion, the estimate of the policy's effect in a pre-expansion year should be $\delta_{DD}^t = 0$ for all years prior to 1993 if the identifying assumptions hold true. That is to say, in the absence of the treatment, no treatment effect should be observed, and the outcomes should trend in parallel. Thus, while the identifying assumption of model (3) cannot ever be tested directly, since they pertain to the unobserved counterfactual outcomes of respondents, they may be tested in pre-expansion periods to argue that the trends (or trends of trends) would have continued to hold in the post-implementation era.

Finally, it may be of interest to estimate the joint significance of the estimated treatment effects across outcomes and models. For instance, even if several behavior indexes were not significantly affected by the policy when considered individually, their collective post-policy changes observed in the data may have been highly unlikely under the assumption that the policy had a zero effect on all scales. To give a specific case, one may wish to test the null hypothesis that each of the difference-in-differences treatment effects of OBRA93 on each of the trichotomous scales was zero:

$$\{H_0 : \delta_{DD}^{BPI\ (Tri)} = \delta_{DD}^{Externalizing} = \delta_{DD}^{Internalizing} = 0\}$$

Similarly, one may also note that a large number of regressions are estimated in this paper. Even in the case of a true null hypothesis (that there is no treatment effect), I could still inadvertently report spuriously significant effect estimates roughly 5% of the time if I reject at the $p \leq 0.05$ level, provided each estimation is independent of the last. In this way, my probability of having committed a Type 1 error could be compounded with each new regression. This possibility underscores the need to conduct a single test (or at least, a reduced number of hypothesis tests) to simultaneously determine the collective significance of the causally-motivated coefficients that I estimate for each of my outcome variables.

The matter is further complicated, however, by the fact that many of the mental health and behavioral problems under consideration in my study are *not* independent of each other. Rather, they exhibit a high degree of co-morbidity, such that individuals with one disorder are at increased risk of exhibiting another. For example, this is certainly true of the incidence of peer conflict and antisocial tendencies; the two outcomes have a positive correlation coefficient of approximately 0.545 (significant at the 0.001 level) in my sample.

To conduct inference across outcomes while accounting for these interdependencies in the data, I first note that any model of each non-cognitive outcome $k$ may be rewritten in matrix notation as

$$Y^k = X^k \beta^k + \varepsilon^k$$

Note that here the superscript $k$ denotes the outcome to which the equation pertains and not an exponentiated power.

From there, I can construct a large stacked regression using a block-diagonal matrix to yield the "seemingly unrelated" affine system

$$\begin{bmatrix} Y^1 \\ \vdots \\ Y^n \end{bmatrix} = \begin{bmatrix} X^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & X^n \end{bmatrix} \begin{bmatrix} \beta^1 \\ \vdots \\ \beta^n \end{bmatrix} + \begin{bmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^n \end{bmatrix}$$

Which may be expressed more compactly as the equation

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathcal{E} \tag{5}$$

The coefficient point estimates contained in $\mathbf{B}$ will be the same as those of the $\beta_k$ generated by each individual regression if the stacked regression model is estimated under standard OLS conditions, such that $\widehat{\mathbf{B}} = (\hat{\beta}_1, \dots, \hat{\beta}_n)^T$ in this case. However, as previously noted, there is good reason to believe that standard Gauss-Markov assumptions are violated. Even if the estimates themselves are unbiased,[4] the outcomes are likely related to one another, as are clusters of observations among individuals and siblings within outcomes, and there is little reason to believe that residual variance is identical for each of the models in the stack. By once again clustering the standard errors of (5) at the household level, I may instead obtain an asymptotically estimated variance-covariance matrix that allows for arbitrary correlations between household members across time periods and outcomes. If child $i$ exhibits a high Anxiety/Depression score in year $t$, while child $i'$ exhibits a high degree of immature emotional dependence in year $t'$, then the potentially nonzero covariance between the residuals of the observations should be accounted for in post-estimation hypothesis testing, provided both children are both the progeny of the same mother $j$.

I may then estimate equation (5) by stacking regressions of the same index type and identification strategy.[5] From there, the joint significance of multiple parameters in $\mathbf{B}$ may

[4]As usual, $\widehat{\mathbf{B}}$ will be unbiased if $\mathbb{E}[\mathcal{E}] = \mathbb{E}[\mathcal{E}|\mathbf{X}] = 0$ and the linear model is correctly specified, even if the other Gauss-Markov assumptions fail.

[5]This means that the stacked model used in this paper stacks regressions that have the same independent variables but different independent variables, such that $X_1 = \cdots = X_n = X_k$ within each $\mathbf{X}$. In general, the stacked regression strongly resembles the Seemingly Unrelated Regression (SUR) model developed in Zellner (1962). In this special case, where the independent regressor variables are the same for each regression,

be evaluated in a manner similar to using a Wald test over the stacked model (Weesie 1999). For this paper, this entails determining the joint significance of all difference-in-differences trichotomous effect estimates and the joint significance of all difference-in-differences dichotomous effect estimates.

## 5. Main Regression Results

After estimating the two-period DID closed-form models specified by equation (3), I report the difference-in-differences estimators of the OBRA93 expansion's effect on child mental health outcomes in Table 3. Each coefficient in Table 3 is generated using a separate regression; the effect on each outcome is estimated once using the unadjusted DID strategy, and then again using the regression-adjusted DID and DDD strategies, from left to right. The DID coefficients represent the change in differences in outcomes between low-income children from single and multi-child households following the passage of the EITC expansion in 1993. Similarly, the DDD coefficients represent this change in differences between the low-income household groups before and after the policy, but with the similarly constructed change in differences between high-income household groups subtracted out, as these latter households were likely unaffected by the policy change. Both methods should, in theory, identify the causal effect (1) of the EITC expansion on the outcome of interest if their assumptions hold. The adjusted models in the latter two columns contain a full set of demographic and socioeconomic controls, while the unadjusted models detailed in the first column do not.

The two identification strategies yield slightly different results for many of the outcomes. For instance, the improvements to the general behavior problems index are significant at the 0.05 level when estimated using the DDD model, while the coefficients are not statistically significant when calculated using difference-in-differences strategies. Even so, the policy's estimated effect on the externalizing behavior subindex is consistently negative and significant across the simple DID, adjusted DID, and adjusted DDD models at the 0.10, 0.05, and 0.01 levels, respectively, suggesting that externalizing behavior was reduced by the increase in household financial resources. Additionally, post-policy reductions in headstrong oppositional behavior are significant at the $\alpha = 0.10$ significance level in both the baseline DID and adjusted DDD models. Other scattered coefficients of related outcomes are also significant at the 0.10 level, like those pertaining to peer conflict and antisocial behavior. None of the other reductions in behavior problems detected by the DID are significantly different from zero.

Yet although only a few of the estimated treatment effects are statistically significant, the vast majority of the coefficients generated by all three sets of regressions in Table 3 are uniformly negative; only those that pertain to emotional dependency and the internalizing behavior baseline DID are an exception to this general pattern. As such, the number of affirmatively answered behavioral problems questions fell faster for the treatment group than the control for most outcomes after OBRA93, suggesting that the treatment group exhibited, on average, fewer behavioral problems than those who did not receive the expanded benefit.

generalized least squares estimation of the classical SUR model with the specification of (5) will also yield a **B** vector equal to a stacked vector of the individual OLS regression estimates (Amemiya 1985).

**Table 3:** DID Estimates of the OBRA EITC Expansion's ITT Effect on Child Outcomes, Years 1990 to 1998, Simple Pre-Post Formulation

| Z-Score of Outcomes | Index Generation Type | (1) DID (Basic) | (2) DID (With Controls) |
|---|---|---|---|
| BPI (Tri) | Trichotomous | -0.0773 | -0.108 |
| | | (0.0676) | (0.0693) |
| External | Trichotomous | $-0.130^{+}$ | $-0.149^{*}$ |
| | | (0.0681) | (0.0704) |
| Internal | Trichotomous | 0.0118 | -0.0533 |
| | | (0.0673) | (0.0759) |
| *Joint P-Value of Estimates Across Outcomes:* | | *0.00540* | *0.0258* |
| BPI | Dichotomous | -0.0651 | -0.0605 |
| | | (0.0665) | (0.0676) |
| Anx/Dep | Dichotomous | -0.0689 | -0.0489 |
| | | (0.0674) | (0.0775) |
| Dependent | Dichotomous | 0.0667 | 0.0857 |
| | | (0.0659) | (0.0738) |
| Hyperactive | Dichotomous | -0.0774 | -0.0945 |
| | | (0.0662) | (0.0697) |
| Headstrong | Dichotomous | $-0.125^{+}$ | -0.0714 |
| | | (0.0669) | (0.0756) |
| Antisocial | Dichotomous | -0.0798 | -0.0832 |
| | | (0.0661) | (0.0701) |
| Peer Conflict | Dichotomous | -0.0352 | $-0.134^{+}$ |
| | | (0.0678) | (0.0800) |
| *Joint P-Value of Estimates Across Outcomes:* | | *0.0891* | *0.121* |
| Observations | | 8202 | 8202 |

Heteroskedasticity-robust standard errors reported in parentheses. Standard errors are clustered at the parent level to allow for arbitrary correlation between members of the same household over time. Each independent variable has been z-scored prior to regression; estimates are in terms of standard deviations. The unadjusted model includes no additional controls, while the regression-adjusted models include region, mother, and year fixed effects, along with dummy variables for child age, parent age, number of siblings, mother's degree of education, and maternal marital status. Additional linear controls in adjusted models include the mother's weekly hours worked, hourly pay, the annual pre-tax wage income of the household, and the total unemployment, AFDC, SSI and foodstamp benefits received by the household in the survey year. The italicized p-values pertain to a joint null hypothesis of zero effect for all outcomes in the given column.

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Given the large number of negative effect estimates in each column of the table, I then attempt to determine the probability of generating a set of coefficients as or more extreme as the ones yielded by the sample, provided the true value of each parameter is zero (such that OBRA93 affected none of the outcomes of interest). This entails using a method analogous to a Wald test over a “seemingly unrelated” stacked regression model, as discussed earlier in this paper. I group the regressions by model type and index type; the first such null hypothesis that I test is as follows:

$$\{H_0 : \delta_{DD}^{BPI\ (Tri)} = \delta_{DD}^{Externalizing} = \delta_{DD}^{Internalizing} = 0\}$$

The p-values resulting from each joint null hypothesis test are reported in italics in Table 3, below the group of coefficients to which they pertain. In aggregate, the coefficients of the three trichotomously generated indexes are *jointly* statistically significant at varying levels for all three of the identification strategies. The negative effect estimates on the dichotomous indexes would also have been, collectively, arguably unlikely in the event that OBRA93 had absolutely no influence on child behavior or mental health; the joint p-values of both the adjusted and baseline DID estimates are near the 0.10 level, although the triple-differences p-value is substantially higher (and far too high to reject a zero-effect null hypothesis, under standard thresholds).

As a final comment on Table 3, it is worth noting that the triple-differences regressions yielded estimates that were greater in both magnitude and statistical significance than their double-differences counterparts. The statistically significant DID coefficients suggest that behavioral problems were generally between one and two tenths of a standard deviation, while the statistically significant triple-differences estimates indicated index reductions of two to three tenths.

Following the calculation of these average post-treatment estimates, I then estimate 4, the extended DID multi-factor model for the trichotomous outcomes; the results are displayed in Table 4. The first row of coefficients in the table denote the “treatment effect” estimates of OBRA93 on each outcome in 1992, the year before the policy was passed. Each pre-policy difference-in-differences is small in comparison to post-treatment years, although they are all negative. Still, none of them are significantly different from zero at even the 0.10 level. As such, the parallel trends identifying assumption of the DID framework does not appear to be implausible for the trichotomous indexes, at least up until the policy was enacted. At the very least, a null hypothesis of no differences in controlled trends fails to be rejected with 90% confidence.

Of the DID estimates in Table 4, only the reductions pertaining to externalizing behavior are significant, first at the 0.05 level in 1996 and then at the 0.10 level in 1998. The second-to-last row of the table reports the p-value of a Wald Test pertaining to the joint significance of the 1996 and 1998 treatment estimates; the estimated effects on externalizing behavior here induce the rejection of $H_0 : \{\delta_{DD}^{1996} = 0,\ \delta_{DD}^{1998} = 0\}$ at just below the $\alpha = 0.10$ level.

The regression-adjusted extended DID estimates of the OBRA93 expansion on the dichotomous BPI and its sub-indexes are similarly reported in Table 5. For these indexes, none of the individual treatment effect estimates in the post-expansion years of 1996 and 1998 are statistically significant at the 0.05 level, although all except for those pertaining to emotional dependency were negative. Additionally, the null hypothesis $H_0 : \{\delta_{DD}^{1996} = 0,\ \delta_{DD}^{1998} = 0\}$ can-

**Table 4:** Estimates of the OBRA93 EITC Expansion's ITT Effect in Each Year on Trichotomous Outcome Indexes, With Full Controls

| | (1) BPI | (2) External | (3) Internal |
|---|---|---|---|
| Effect in 1992 | -0.0323 | -0.0275 | -0.0764 |
| (Trends Check) | (0.0625) | (0.0620) | (0.0719) |
| Effect in 1996 | -0.128 | -0.169* | -0.0655 |
| | (0.0803) | (0.0813) | (0.0930) |
| Effect in 1998 | -0.133 | -0.168 | -0.155 |
| | (0.0959) | (0.0960) | (0.103) |
| Joint P-Value) | 0.250 | 0.0994 | 0.308 |
| Obs. | 8202 | 8202 | 8202 |

Heteroskedasticity-robust standard errors reported in parentheses. Standard errors are clustered at the parent level to allow for arbitrary correlation between members of the same household over time. Each independent variable has been z-scored prior to regression; estimates are in terms of standard deviations.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

not be rejected for any of the dichotomously measured outcomes at the standard levels of significance.

## 6. Discussion

In aggregate, there does appear to be some empirical evidence that the OBRA93 EITC expansion reduced the incidence of affective disorders among children and adolescents, as reported by their parents in the behavioral problems indexes of the NLSY Child supplement. This was particularly true for measures of externalizing behavior; the estimated effect of the policy change on the externalizing behavior subindex was consistently negative and significant across models. Additionally, several of the smaller dichotomous subindexes related to externalizing behavior, such as those pertaining to headstrong oppositional defiance, antisocial tendencies, and peer conflict, were significant at varying levels under different strategies as well.

On a more general level, nearly all of all of the DID treatment effect estimates calculated in this paper were negative; many of these improvements in outcomes were jointly significant when grouped across models of the same type. This suggests that the number of behavioral problems reported by the treatment group fell faster than the number reported by the control group in the years that followed the OBRA93 EITC expansion, although many of these differences were not independently statistically significant. Furthermore, the parallel trends assumption of the difference-in-differences identification strategy seemed plausible for the majority of the outcomes of interest – including those related to externalizing behavior – given the graphical and statistical inspection of the assumption in pre-treatment years.

**Table 5:** DID Estimates of the OBRA93 EITC Expansion's ITT Effect in Each Year on Dichotomous Outcome Indexes, Full Controls

| | (1) BPI | (2) Anx/Dep | (3) Dependency | (4) Hyperactivity | (5) Headstrong | (6) Antisocial | (7) Peer Conflict |
|---|---|---|---|---|---|---|---|
| Effect in 1992 | -0.0319 | -0.0732 | -0.0516 | -0.0943 | 0.0599 | 0.00687 | 0.0289 |
| (Trends Check) | (0.0614) | (0.0721) | (0.0761) | (0.0736) | (0.0694) | (0.0623) | (0.0786) |
| Effect in 1996 | -0.0793 | -0.0707 | 0.0904 | -0.147 | -0.0827 | -0.0605 | -0.0895 |
| | (0.0803) | (0.0962) | (0.0909) | (0.0882) | (0.0876) | (0.0833) | (0.0990) |
| Effect in 1998 | -0.0835 | -0.133 | 0.0155 | -0.173 | 0.0381 | -0.102 | -0.166 |
| | (0.0916) | (0.101) | (0.105) | (0.0972) | (0.106) | (0.0952) | (0.107) |
| Joint P-Value | 0.573 | 0.414 | 0.502 | 0.162 | 0.222 | 0.558 | 0.297 |
| Obs. | 8202 | 8202 | 8202 | 8202 | 8202 | 8202 | 8202 |

Heteroskedasticity-robust standard errors reported in parentheses. Standard errors are clustered at the parent level to allow for arbitrary correlation between members of the same household over time. Each independent variable has been z-scored prior to regression; estimates are in terms of standard deviations. All models include region, mother, and year fixed effects, along with dummy variables for child age, parent age, number of siblings, mother's degree of education, and maternal marital status. Additional linear controls include the mother's weekly hours worked, hourly pay, the annual pre-tax wage income of the household, and the total unemployment, AFDC, SSI and foodstamp benefits received by the household in the survey year. The joint p-value pertains to the joint significance of the treatment effect estimates in years post-OBRA93.
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Although the triple-differences assumptions were somewhat more questionable, the effect estimates were also similarly negative, suggesting some agreement between models that most of the behavior problems metrics were lower with the treatment on average.

But in the case that the the main identifying assumptions of the DID and DDD strategies are valid, endogenous selection of households from the single-child control group into the multi-child treatment group may have occurred, as household birth rates did increase differentially following 1993. While this did not appear to significantly affect estimates of the policy's treatment effect on child affective disorders, especially since the models included household fixed effects and controls for variables such as household size, it still represents a possible confounder in the evaluation of the policy's impact.

One may also note that the differential expansion of the EITC – the treatment delivered by OBRA93 – was limited to approximately \$1400 dollars a year or less; even in 1993, this was a modest, albeit not insignificant, sum. It is possible that a larger positive income shock could have changed the behavior and expenditures of recipient households more significantly, which in turn could then have affected the well-being of low-income children more strongly. Of course, more research on larger expansions is necessary to more thoroughly test this hypothesis. The returns of yearly financial transfers to child mental health could compound with time; many of the effect estimates reported by the extended multifactor models in 1998 are greater in magnitude than those of 1996, although the coefficients lack the precision to state this conclusively.

There are other reasons to believe that significant cash transfers to low-income households could affect long-term child mental health outcomes to a greater extent than detailed here. This paper largely examines the outcomes of children before and after the policy's implementation. But poverty may have lasting effects on individuals who grow up in it that cannot be so easily remedied by assistance after-the-fact. My study may thus understate the preventative efficacy of transfers made to children before or during their early formative years. More concretely, financial assistance could reduce exposure to risk factors correlated with impoverishment that cause long-lasting or even irreversible harm. This could be especially true with environmental factors like lead exposure in adolescence. For instance, a large literature in neuroscience and environmental policy has found substantial evidence that lead exposure in early childhood is strongly linked with increased rates of mental disorders and violent crime later in life, along with decreased intelligence and cognitive function (Bellinger 2008). A financial transfer cannot undo this damage. However, if of sufficient magnitude, the pecuniary assistance could prevent household exposure by reducing the rate at which low-income families are priced into areas of high lead concentration, which could in turn lower the lifelong harm to children and adolescents. These potential lasting connections between financial resources and mental health are largely beyond the scope of this study, despite their potential importance.

Despite the limitations of the methodology of this paper, my central finding that externalized behavior is reduced by expanded household resources generally corroborates the conclusions of Milligan and Stabile (2011), who found that higher transfers reduced physical aggressiveness in children. While their results also suggested that children on average tended to exhibit fewer beneficent "pro-social" behaviors (like helping one another and volunteering) after their families received a higher pecuniary benefit, reductions in aggressiveness would be congruous with reduced externalized behavioral problems in general. My estimates and those of Milligan and Stabile (2011) are also both of roughly the same order of magnitude: between roughly $-0.10$ and $-0.25$ standard deviations.

This paper might also complement the findings of Dahl and Lochner (2012), who found that the EITC improved academic performance; children who act out less in general might also be less disruptive in their classroom, which could lead to increases in learning. Lastly, behavioral problems can extend well beyond childhood; externalized behavior in youth can often predict violence and crime later in life (Liu 2006). As such, even if the effect of transfer policies like the EITC on other types of child behavioral problems and affective disorders are somewhat muted, reductions to externalized behavior could still yield significant returns to social welfare.

In all, I find that the federal Omnibus Budget Reconciliation Act of 1993's expansion of the earned income tax credit reduced the prevalence of some affective mental health disorders and behavioral problems among children from low-income households. In particular, the policy appeared to decrease the incidence of externalizing behavior. Yet while the main externalizing behavior results were comparatively robust, many of the other intent-to-treat effects pertaining to other more specific behavior problem sub-indexes were barely significant at the 0.10 and 0.05 levels, if at all. Additionally, the potential responsiveness of households to the EITC, in terms of working more hours and having more children, could compromise identification within my difference-in-differences framework. As such, while the general body of evidence suggests that federal cash transfer programs like the EITC may positively influence the mental health of children, my findings are not wholly unambiguous.

## 7. Bibliography


Amemiya, Takeshi (1985). Advanced Econometrics. Cambridge, Massachusetts: Harvard University Press.

Angrist, Joshua D., and Jörn-Steffen Pischke (2009). Mostly Harmless Econometrics. Princeton, NJ; Princeton University Press.

Averett, Susan and Yang Wang (2015). The effects of the earned income tax credit on children's health, quality of home environment, and non-cognitive skills. IZA Discussion Paper No. 9173.

Bellinger DC (2008) Neurological and Behavioral Consequences of Childhood Lead Exposure. PLoS Med 5(5): e115. doi:10.1371/journal.pmed.0050115.

Baughman, Reagan and Stacy Dickert-Conlin (2003). Did Expanding the EITC Promote Motherhood? American Economic Review Papers and Proceedings 93(2): 247-250.

Baughman, Reagan and Stacy Dickert-Conlin (2009). The Earned Income Tax Credit and Fertility. Population Economics 22(3): 537-563.

Bertrand, Marianne, Esther Duflo, Sendhil Mullainathan (2004). How Much Should We Trust Differences-In-Differences Estimates?. Quarterly Journal of Economics 2004; 119 (1): 249-275.

Boyd-Swan, C., Herbst, C. M., Ifcher, J., & Zarghamee, H. (2013). The earned income tax credit, health, and happiness. Discussion Paper No. 7261, Institute for the Study of Labor (IZA).

Cancian, M., Levinson, A. (2005). Labor Supply Effects of the Earned Income Tax Credit: Evidence from Wisconsin's Supplemental Benefit for Families with Three Children. NBER Working Paper No. 11454.

Center on the Developing Child (2013). Early Childhood Mental Health (InBrief). Retrieved from www.developingchild.harvard.edu.

Crump, R., Goda, G., and Mumford, K. (2011). Fertility and the Personal Exemption: Comment. The American Economic Review, 101(4), 1616-1628.

Cunha, F., J. J. Heckman, L. J. Lochner, and D. V. Masterov (2006). Interpreting the Evidence on Life Cycle Skill Formation. In E. A. Hanushek and F. Welch (Eds.), Handbook of the Economics of Education. Amsterdam: North-Holland.

Dahl, Gordon B. and Lance Lochner (2012). The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. American Economic Review, 102(5): 1927-56.

Donald S. Kenkel, Maximilian D. Schmeiser & Carly Urban (2014). Is Smoking Inferior?: Evidence from Variation in the Earned Income Tax Credit," Journal of Human Resources, University of Wisconsin Press, vol. 49(4), pages 1094-1120.

Eissa, Nada and Hilary W. Hoynes, 2006. "Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply," NBER Chapters in: Tax Policy and the Economy, Volume 20, pages 73-110 National Bureau of Economic Research, Inc.

Evans, W., and C. Garthwaite (2011). Giving Mom a break: The effect of higher EITC payments on maternal health. Working paper 18206, National Bureau of Economic Research.

Goetzel, Ron Z, Long, Stacey R., Ozminkowski, et al. (2004). Health, Absence, Disability, and Presenteeism Cost Estimates of Certain Physical and Mental Health Conditions

Affecting U.S. Employers. *Journal of Occupational and Environmental Medicine*, Volume 46, Issue 4, pages 398-412.

Hamad, R., Rehkopf, D.H. (2016). Poverty and Child Development: A Longitudinal Study of the Impact of the Earned Income Tax Credit. American Journal of Epidemiology.

Hotz, V. Joseph and John Karl Scholz (2003). The Earned Income Tax Credit; Means-Tested Transfer Programs in the United States. Edited by Robert A. Moffitt. University of Chicago Press.

IRS (2016). Calendar Year Report. The Internal Revenue Service of the United States.

Kim, DaeHwan and J. Paul Leigh (2011). Are Meals at Full-Service and Fast Food Restaurants "Normal" or "Inferior"? Population Health Management. December 2011, 14(6): 307-315.

Lino, Mark (1997). Expenditures on Children by Families, 1996 Annual Report. U.S. Department of Agriculture, Center for Nutrition Policy and Promotion. Miscellaneous Publication No. 1528-1996.

Liu, J. (2004). Childhood Externalizing Behavior: Theory and Implications. Journal of Child and Adolescent Psychiatric Nursing: Official Publication of the Association of Child and Adolescent Psychiatric Nurses, Inc, 17(3), 93103.

Maxfield, M. (2013). The Effects of the Earned Income Tax Credit on Child Achievement and Long-Term Educational Attainment. Michigan State University Job Market Paper.

Maxfield, M. (2014). The Effects of the Earned Income Tax Credit on Net Family Financial Resources. Michigan State University.

McLeod, J., and Shanahan, M. (1993). Poverty, Parenting, and Children's Mental Health. American Sociological Review, 58(3), 351-366.

Mensah, F.K., and Kiernan K.E. (2010). Parents' Mental Health and Children's Cognitive and Social Development; Families in England in the Millennium Cohort Study. Social Psychiatry and Psychiatric Epidemiology, Volume 45: 1023-1035.

Milligan, Kevin and Mark Stabile (2011). Do Child Tax Benefits Affect the Well-Being of Children? Evidence from Canadian Child Benefit Expansions. American Economic Journal: Economic Policy, 3(3): 175-205.

Milligan, Kevin (2005). Subsidizing the Stork: New Evidence on Tax Incentives and Fertility. The Review of Economics and Statistics, MIT Press, vol. 87(3), pages 539-555, 06.

Mills P. R., Kessler R. C., Cooper J., Sullivan S (2007). Impact of a health promotion program on employee health risks and work productivity. American Journal of Health Promotion. 2007;22:4553.

Nemets, Hanah, Boris Nemets, Alan Apter, Ziva Bracha, and R.H. Belmaker (2006). Omega-3 Treatment of Childhood Depression: A Controlled, Double-Blind Pilot Study. *The American Journal of Psychiatry*, June 2006, Vol 163 No 6.

O'Neil, A., Quirk, S. E., Housden, S., Brennan, S. L., Williams, L. J., Pasco, J. A., et al. (2014). Relationship Between Diet and Mental Health in Children and Adolescents: A Systematic Review. American Journal of Public Health, 104(10), e31-e42.

Pischke, Jörn-Steffen (2005). Empirical Methods in Applied Economics: Lecture Notes. The London School of Economics.

Pratt, Laura and Debra J. Brody (2014). Depression in the U.S. Household Population, 2009-2012. National Center for Health Statistics, Data Brief No. 172.

Proctor, Bernadette D., Jessica L. Semega, and Melissa A. Kollar (2016). Income and Poverty in the United States: 2015. The United States Census Bureau, Report Number P60-256.

Smith, J. P., Monica, S., and Smith, G. C. (2010). Long-Term Economic Costs of Psychological Problems During Childhood. Social Science & Medicine (1982), 71(1), 110115.

Weesie, J. (1999). SG121: Seemingly unrelated estimation and the cluster-adjusted sandwich estimator. Stata Technical Bulletin 52: 3447. Reprinted in Stata Technical Bulletin Reprints, vol. 9, pp. 231248. College Station, TX: Stata Press.

Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data. Cambridge, Mass: MIT Press.

Zellner, Arnold (1962). An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. Journal of the American Statistical Association.

## 8. Appendix

### *8.1. Appendix A: Labor Supply Response*

In this section, I attempt to determine the average effect of the Omnibus Reconciliation Act's expansion of the earned income tax credit on the labor supply outcomes of women with children. To do so, I once more exploit the differential expansion of $OBRA93$, in which multi-child households received a larger expansion than single-child ones. I thus construct a difference-in-differences model using likely EITC eligible mothers with two or more children as a treatment group and mothers with only one child as a control group. My dataset remains the NLSY79 over the years 1990 to 1998, where likely EITC eligible mothers are defined as women with children who made less than or equal to $30,000 in 1992.

In general, my DID findings are consistent with those found by Evans and Garthwaite (2011) and the general literature: recipients of the credit increased their labor supply on both the intensive and extensive margin. Women with children responded to the EITC by working nearly 2 more hours per week on average, while the percentage of low-income mothers who were employed similarly rose by over 5%.

**Table 6:** DID Estimates of the ITT Effect of the OBRA EITC Expansion on Work Outcomes of Women with Children, Years 1990 to 1998

| Labor Supply Outcome | All | Unmarried | Married |
|---|---|---|---|
| Hours Worked | $1.867^{*}$ | $1.739^{+}$ | 1.151 |
| | (0.780) | (1.035) | (1.232) |
| Percentage Employed | $0.0559^{*}$ | 0.0341 | $0.0647^{+}$ |
| | (0.0222) | (0.0298) | (0.0367) |
| Observations | 6,399 | 3,578 | 2,821 |

Heteroskedasticity-robust standard errors reported in parentheses. Standard errors are clustered at the respondent level to adjust for serial correlation across time periods. All models include region and year fixed effects and dummy variables for the respondent's age, race, number of children, level of education, and marital status. Additional linear controls include the respondent's hourly pay, the annual pre-tax wage income of both the respondent and their spouse, and the total unemployment, AFDC, SSI and foodstamp benefits received by the household.

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$