CARRA Working Paper Series

Working Paper *#2014-13*

Do Doubled-up Families Minimize Household-level Tax Burden?

Maggie R. Jones
Amy B. O'Hara
U. S. Census Bureau

Center for Administrative Records Research and Applications
U.S. Census Bureau
Washington, D.C. 20233

Paper Issued: September 14, 2014

Disclaimer: This paper is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

Do Doubled-up Families Minimize Household-level Tax Burden[1]?

Maggie R. Jones, *U.S. Census Bureau*
Amy B. O'Hara, *U.S. Census Bureau*

September 17, 2014

Abstract

This paper examines a method of tax avoidance not previously studied: the sorting of dependent children among related filers who have "doubled up" in a household for economic reasons. Using the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) linked with 1040 data from the Internal Revenue Service (IRS), we examine households with children and at least two adult tax filers to determine whether the household minimizes income tax burden, and thus maximizes refunds, by optimally claiming dependents. We examine specifically the relationship between the Earned Income Tax Credit (EITC) and the sorting of dependent children among filers in households. We find the following: The propensity to sort increases as the number of filers who are potentially eligible for the EITC increases; sorting probability increases as the optimal household EITC amount increases; and among households with at least one EITC-eligible filer, the propensity to sort increases as the difference between modeled household EITC amount and the optimal amount increases. We also exploit the 2009 change in EITC benefit for families with three or more children, finding that the propensity to sort to exactly three children increased among EITC-eligible filers after the rule change. The results of this analysis improve our understanding of filing behavior, particularly how households form filing units and pool resources, and have implications for poverty measurement in complex households.

[1] The views expressed in this paper are those of the authors and do not necessarily represent the views of the U.S. Census Bureau.

1. INTRODUCTION

Recently, there has been an increasing focus in the labor economics literature on behavioral responses to discontinuities in the tax code.[2] This literature is not so much focused on responses to an overall tax policy, but rather the way in which tax filers take advantage of the mechanical quirks of the tax system in order to gain as much out of the system as possible. Most of the activity analyzed, such as the bunching of income of self-employed taxpayers at kink points in the Earned Income Tax Credit's (EITC) benefit structure, or the claiming of children who do not exist, appears to be outside the letter of the law. However, as with any set of rules as complicated as the U.S. federal income tax system, there exist interpretations and quirks that may be exploited in a legal manner. We examine such behavior in this paper, which looks at how multiple adult tax filers in a household are able to sort all the dependent children in the household in such a way as to minimize overall household tax burden. The behavior we examine falls into a specific category of tax avoidance—tax arbitrage between members of a family (Stiglitz, 1988). To our knowledge, ours is the first paper to look at the issue of optimal claiming of dependents by related household members. Our access to a unique data set—linked survey and 1040 data—gives us the necessary information to examine the issue.

There are several reasons why the question of dependent sorting is important. As with all tax avoidance, there are implications for public finance and tax-system equity (Andreoni, Erard, & Feinstein, 1998). The question of how child dependents are sorted in multifamily households is an important one for assessing their well-being. We show that the difference in household tax burden or refund can be as large as $4,000 (for the 2010 tax year) depending on the choice of which filers in a household claim the dependents. Moreover, because of the EITC benefit structure, differences tend to be higher for households where at least one filer has low income.

[2] See, for example, Saez (2010) and LaLumia and Sallee (2012).

On the one hand, flexibility in tax rules may be an important policy goal, since it provides low-income, resource-pooling households with filing options that increase the total household refund. On the other hand, such flexibility may backfire if a filer does not actually take responsibility for the dependents he or she claims, since any refund money might go solely to the person claiming it rather than the household (for an example of such a concern playing out in terms of pension payments, see Duflo (2003)).

The question of who claims dependents also has implications regarding our measurement of complex households. For example, the alternative poverty measure calculated by the U.S. Census Bureau relies on precise estimation of transfers from the tax system and depends on an understanding of household structure and which adults in a household might be financially responsible for which children. Doubling up has long been used by young families in response to high rents or low wages (Haurin, et al., 1993). During the Great Recession, "doubling-up" became a more common strategy for making ends meet, especially among low-income households that rely on tax credits as part of the social safety net (Mykyta & Maccartney, 2011).

Finally, the sorting of dependents also gives insight into how well tax rules are understood, either by taxpayers or tax preparers, as well as how that information might be disseminated from household to household. Several recent papers have looked at this question in regards to the EITC specifically (Chetty, Friedman and Saez, 2013; Chetty and Saez , 2009).

The paper proceeds as follows. First, because this is the first paper to look at the question of dependent sorting, in Section 2 we describe at length the various credits and deductions in the tax code related to dependent children, the specific definition of sorting we use in the paper, and some considerations about where to place this behavior on the spectrum of tax avoidance/evasion. We present some evidence from simulations using synthetic tax data showing

how families might minimize their overall tax burden. We also cover some of the relevant literature on tax avoidance. Section 3 describes the data we use for the analysis: the Current Population Annual Social and Economic Supplement (CPS ASEC) linked person-by-person to Internal Revenue Service data for tax years 2005–2010. Section 4 describes the empirical model we use to examine dependent sorting, and Section 5 presents the results of our analysis. Section 6 concludes.

2. BACKGROUND AND LITERATURE

2.1. Deductions and credits in the tax code

Holding everything else constant, a taxpayer is always as well off or better off if he or she is able to claim a dependent (Ellwood & Liebman, 2001), although the value of claiming a dependent varies over filers' incomes. The dependent exemption lowers taxable income for any taxpayer claiming it, so its value depends on the given tax bracket. For taxpayers whose dependents are children, the Child Tax Credit (CTC) and the EITC often apply. Head of household filing status, which may be used by otherwise single filers when a qualifying dependent can be claimed, allows filers a higher standard deduction and wider tax brackets. Each of these deductions and credits have rules that overlap but that are not perfectly coincident. The dependent exemption and head of household status have the same rules governing qualifying children, described below. EITC rules and CTC rules are each slightly different from the dependent exemption.

For the dependent exemption, a qualifying child must meet relationship, age, residency, support, and joint return requirements. The child must be related to the claimant biologically (son, daughter, grandchild, sibling, etc.) or must be a formally adopted or foster child. A qualifying child must be either less than 19 or less than 24 and a full-time student, or

permanently and totally disabled and any age. The child must have lived with the claimant for more than half of the tax year, and the claimant must have provided more than half of the support for the child during that year. Finally, a child cannot have filed a joint return for the same tax year unless the child and spouse filed only to claim a refund (i.e., were not legally required to file).

For head of household status, a filer must be unmarried (or "considered unmarried'") on the last day of the tax year and must have paid more than half of the cost of keeping a home. A filer must also have a qualifying dependent (not necessarily a child) living with them for more than half of the year. For qualifying children, the filer must also claim the child for the dependent exemption; thus, the qualifying child rules for head of household filing are the same as those for the dependent exemption.

The "support rule" for the EITC differs in an important way. While the rules on age and relationship status are the same as those for the dependent exemption, in the case of the EITC, children must live more than half of the year with the claimant, but the support test does not apply. On the other hand, to claim the CTC, a child must be claimed as a dependent on the return, and thus must meet all of the requirements listed for a dependent. Unlike the EITC, the CTC only applies for children younger than 17.

2.2. Sorting of dependents

We define sorting based on household relationship information in the CPS ASEC and the rules regarding the claiming of dependents. For the purposes of our analysis, potential sorting households must have more than one adult 1040 tax filer; we define adults as 18 years of age or older; and we exclude those who are claimed on someone else's return. A household also must

have at least one dependent claimed on a 1040 form by one of the adult filers. We determine that sorting has occurred when the number of children modeled as claimable by the reference person (modeling is based on relationship and income responses in the CPS ASEC) is less than the number claimed in tax records, while simultaneously the number modeled for a second filer in the same household is more than the number claimed. Sorting is also defined in the reverse case, when the number of claimed children is more than the number of modeled children for the reference person and less for the second filer. We are thus exploiting the differences between tax filing behavior based on the survey data and actual behavior reflected in the IRS data. Because of tax rules regarding the relationship of children to those claiming them, we only include relatives of the household head as possible filers to sort to or from. These include children, grandchildren, siblings, parents, and "other relatives" of the CPS reference person.[3]

We are thus examining sorting from the CPS reference person to or from other related filers in the CPS household, but we do not examine sorting between other household members. Because the CPS questionnaire only asks about relationships between the reference person and other members of the household, this choice is due to recognizing the difficulty involved in determining relationships between other household members. For example, in a cohabiting household, we would not be able to tell whether a person who is unrelated to the reference person *is* related to the cohabitor. It should be noted, however, that we capture instances when the reference person sorts to or from multiple household members.

2.3 Simulations showing possible full tax outcomes

[3] We model the reference person as the tax filer in cases where the reference person answered the CPS questionnaire but his or her spouse files the 1040 form for the tax year.

Although an individual filer is always better off if he or she can claim a child, when it comes to sorting children among filers, the goal is to minimize overall household tax burden and maximize household refunds. A household might not achieve an optimal tax outcome through sorting. For example, a household with two children might be better off overall if one person claims both children, even if claiming one of the children would help another member of the household minimize his or her individual tax burden.

To examine the interaction between taxes, credits, filing status, and the benefits of sorting, we created simulated tax data and used NBER's TAXSIM (see Feenberg and Coutts (1993) for more information) to calculate overall tax burden for two cases. In both cases, a single mother with two children lives with her mother, who we call the grandmother. Both women are 1040 filers. The grandmother is also unmarried, and if surveyed in the CPS, would be the reference person. In the first case, the grandmother claims no children and the mother claims both children, and in the second case the grandmother and mother each claim one child. For both mother and grandmother, we allow earnings to range from $1,000 to $100,000, and we use the average values of interest, rental, and Supplemental Security Income (SSI) from our true data (the linked CPS and IRS file) from 2010, rounding to the nearest $100. Using the grandmother as the first filer, we created 100 observations for earnings (in $1,000 increments) for her, and calculated her tax burden as either a single filer with no dependents or a head of household filer claiming one child. For each value of the mother's earnings, we calculated her tax burden as head of household claiming either one or two children. Finally, for each value of the grandmother's earnings, we linked every value for the mother (leading to a simulated data set of 10,000 observations). The difference between the two cases is simply overall household tax burden in case 1 (grandmother's tax plus mother's tax) minus overall household tax burden in

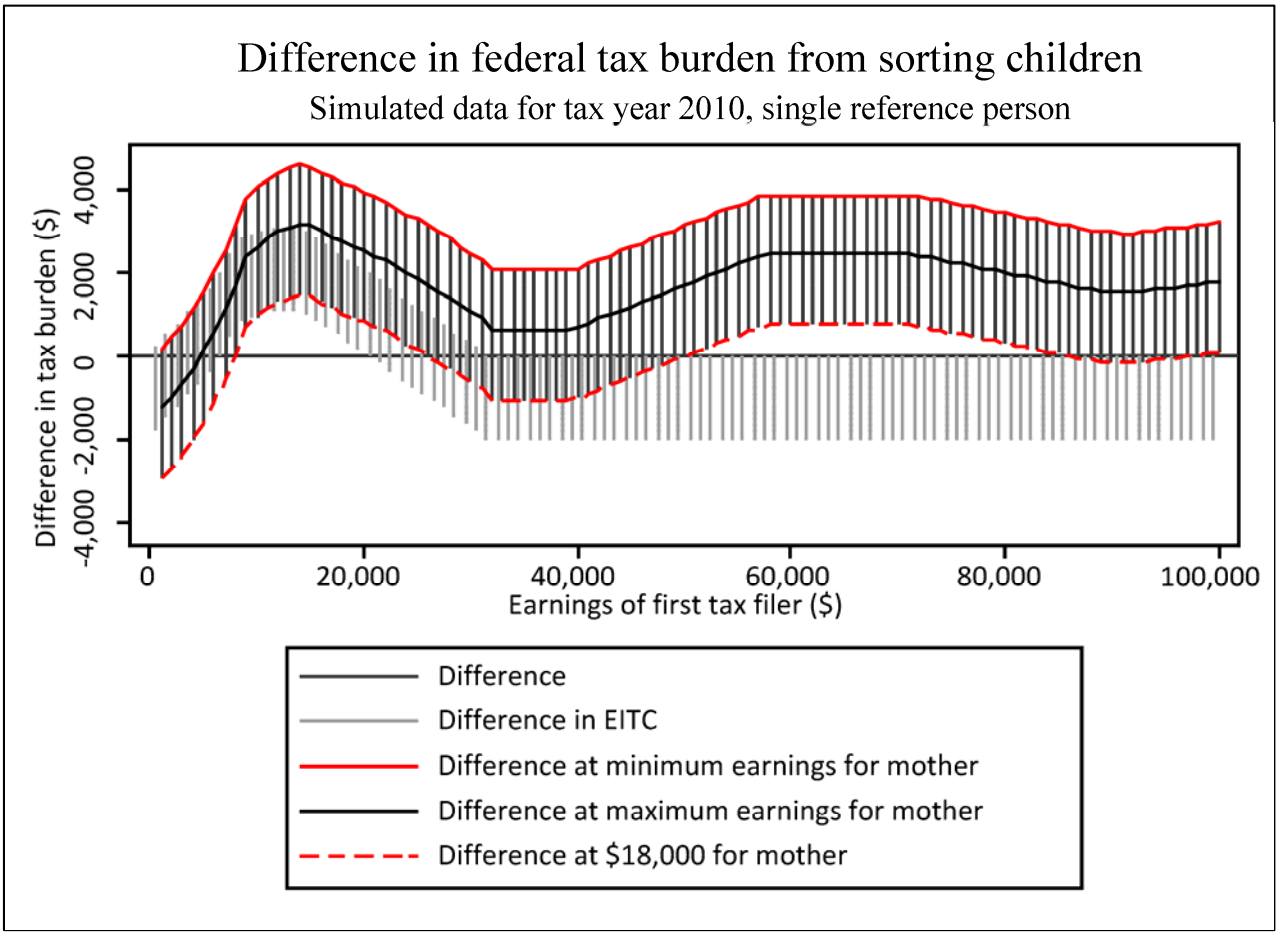

Figure 1. Source: Invented data for tax year 2010, with tax outcomes generated using NBER's TAXSIM program. A difference in tax outcome for a household is calculated for two scenarios: When two dependent children are claimed by their mother, and when one child is claimed by an unmarried grandmother who files singly under the first scenario and head of household under the second. All possible differences are graphed as a function of the grandmother's income (black bars). To show the influence of EITC receipt on the difference, EITC differences are graphed separately in gray (but are included in the total difference).

case 2. The differences in tax burden, including all taxes and credits, are graphed as vertical black bars vis-à-vis the grandmother's earnings in Figure 1. The gray bars show the difference in EITC only. Differences above the *y* axis indicated combinations of earnings for which the family is better off if it sorts.

As can be seen in the figure, for nearly every value of the grandmother's earnings in combination with all possible earnings from the mother, the household described is as well off or better off if it sorts children between the two earners. For situations when the grandmother's earnings are on the extreme low end of the scale, there is a detriment to sorting unless the

mother's earnings are also extremely low. The family would be better off if the mother claimed both children. For a few cases when the grandmother's income is modest (between approximately $30,000 and $50,000) and the mother's is also modest (between $18,000 and $30,000) the family is also better off if they do not sort. Maximum values for sorting occur when the grandmother has low to modest earnings and the mother has very low earnings. For example, the maximum difference of $4,595 is achieved when the grandmother's earnings are $14,000 and the mother's are $1,000 (mainly because the grandmother's earnings place her at the maximum point in the EITC schedule). Values close to the maximum are also achieved when the grandmother's earnings are modest to high and the mother's are low.

The figure demonstrates that, for any value of the grandmother's earnings, a maximum difference occurs when the mother's earnings are at a minimum ($1,000). Because of her low earnings, the gain the mother receives from claiming a second child versus the gain the grandmother receives in claiming one is widest at this point. Meanwhile, for any value of the grandmother's earnings, a maximum value for the mother puts them in an area where the difference in tax burden is average. Interestingly, the same value for the mother's earnings, $18,000, constitutes the point where the difference between the two filing cases is minimized for every choice of the grandmother's earnings.

Difference in total EITC plays a large positive part in the overall difference until the grandmother's earnings are larger than the ending income for the EITC. At that point, it is more beneficial, in terms of EITC, for the mother to claim both children. However, the benefit to the household of both mother and grandmother claiming head of household status is large, and for most combinations of incomes outweighs the loss in EITC. This outcome is different when the grandmother is married and files jointly, as shown in Figure 2.

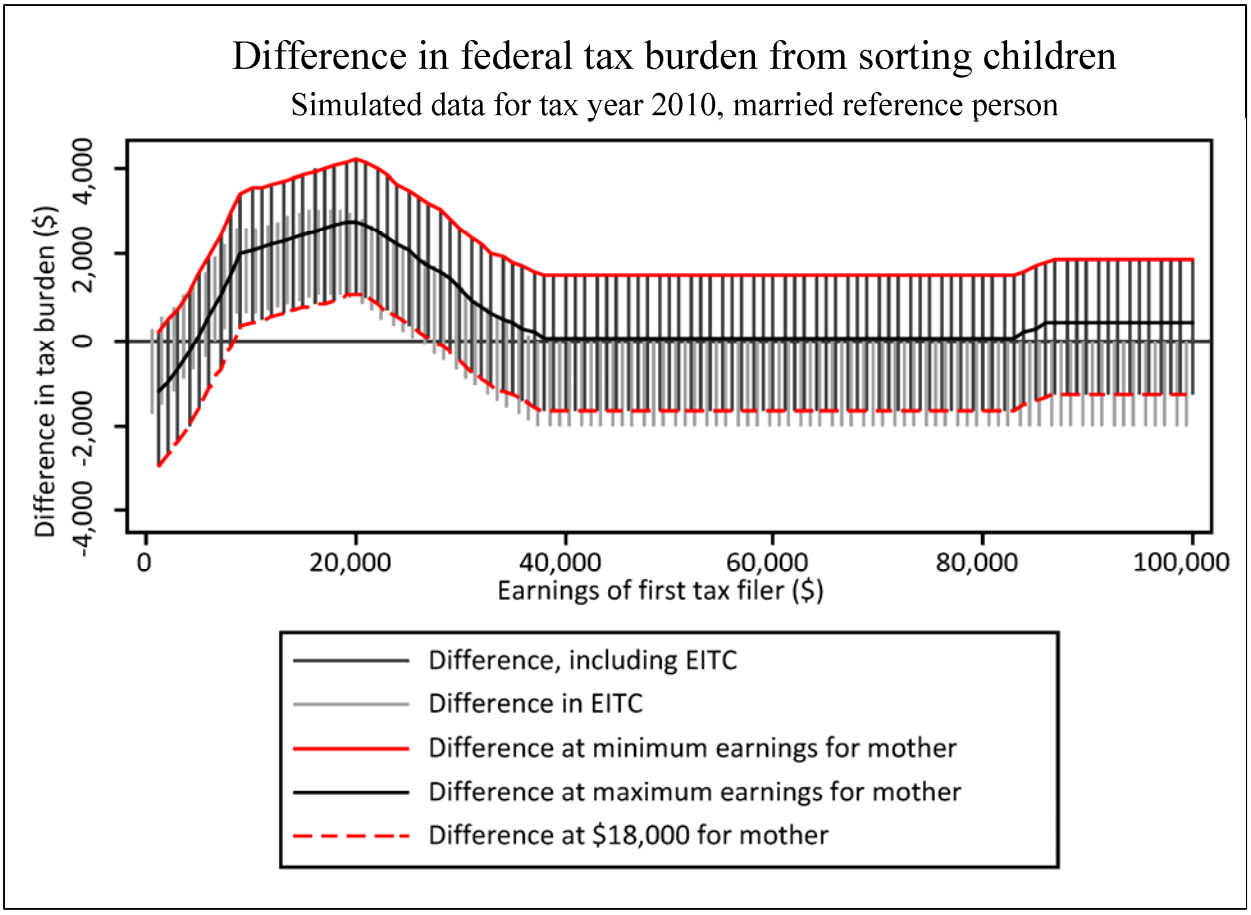

Figure 2. Source: Invented data for tax year 2010, with tax outcomes generated using NBER's TAXSIM program. A difference in tax outcome for a household is calculated for two scenarios: When two dependent children are claimed by their mother, and when one child is claimed by an married grandmother who files jointly under both. All possible differences are graphed as a function of the grandmother's income (black bars). To show the influence of EITC receipt on the difference, EITC differences are graphed separately in gray (but are included in the total difference).

Figure 2 shows that once the grandparent's earnings are above the threshold for the EITC, the family is better off from sorting only about half of the time (and usually when the mother's earnings are very low—$1,000 to $10,000). In this case, the gain the grandparents receive from the adjustments to their taxable income (the dependent deduction and child tax credit) outweigh the small gain the mother gets from the difference in benefit level in EITC and Child Tax Credit from two children versus one child.

2.4. Tax avoidance or evasion?

An important question to ask is whether any observed sorting falls under the definition of tax avoidance or evasion. Slemrod and Yitzhaki (2002) provides definitions of each issue, with illegality being the distinguishing characteristic of evasion. Thus while avoidance is a result of choices that fall within taxation rules, evasion generally boils down to weighing the costs of compliance versus the cost of getting caught. An analysis of avoidance would include the costs of interpreting rules and taking advantage of them, while evasion would include the possibility of an audit and potential penalties as part of its cost structure. It is impossible to know, using the data available, whether the sorting we observe falls under perfectly acceptable interpretations of tax rules or whether it would qualify as rule-breaking when subjected to an audit. For the purposes of this paper, we make the assumption that any sorting of children we observe is allowed by income tax rules.

The issue of tax avoidance—specifically personal income tax avoidance—has received less attention in the economic literature than has tax evasion. When it has been examined, the attention has focused on the accounting problem presented by tax avoidance of various kinds, including the avoidance of sales and corporate taxes (see, for example, Feldstein (1999) and Slemrod (2007), among many others). Andreoni, Erard, & Feinstein (1998) provides a review of the issues surrounding personal income tax compliance, including general equilibrium considerations and auditing rules. In his general theory of income tax avoidance, Stiglitz (1988) outlined and modeled three, possibly overlapping, methods of avoidance: postponement of taxes, tax arbitrage across individuals who face different tax brackets, and tax arbitrage across income streams that face different schemes of taxation. The behavior we model falls firmly into the second category, although Stiglitz does not mention the treatment of dependents for arbitrage. To

the extent that a dependent child represents a tax savings, the transfer of the child to one filer from another constitutes a form of ``tax induced transaction'' that improves a household's tax standing without incurring a monetary cost.

Research into the mechanics of intra-family arbitrage as a response to specific aspects of tax code is scarce. One stream examines transfers between spouses facing individual income taxes. Because individual taxation of spouses occurs mainly in Western European countries, the research is concentrated there. An example is Stephens Jr & Ward-Batts (2004), who found that a change from joint to individual taxation in the UK led to a shift in the share of asset income claimed by wives. Because of the generally lower marginal income tax rate that wives face, households can make a Pareto improvement in their tax position by transferring asset income from husbands to wives. A second stream examines gifts and bequests to family members. For example, Ohlsson (2007) uses a design quirk of the Swedish tax system to investigate how often heirs avoid an inheritance tax by passing on the inheritance immediately to grandchildren. The author found that the propensity to pass down the inheritance increased with the size of the tax.

When dependent claiming has been studied, it has fallen under the category of tax evasion—specifically, the case of nonexistent dependents rather than the case of who claims whom. For example, LaLumia & Sallee (2012) investigated a change in rules in the US income tax code between years 1986 and 1987 that required taxpayers, for the first time, to report Social Security Numbers for dependents on tax returns. The initial rule limited the requirement to children age 5 and older, with the rule applying to increasingly younger dependents in further tax years. The authors found a sharp decrease in the number of dependents reported (a loss of 5.5 percent, equivalent to 4.2 million children) in tax year 1987, an effect attributable to cheating in the preceding tax year.

The value of claiming a child that does not exist varies depending on where a taxpayer falls in the income distribution, their filing status, and the other types of credits he or she might be able to claim (LaLumia & Sallee, 2012). For low-earning taxpayers, the EITC represents a valuable incentive to claim children. Meyer & Rosenbaum (2001) examined the issue of noncompliance and mystery children specifically for the EITC using tax data from 1994, finding that an increase of 10 percent in the EITC benefit was associated with a 4 percent increase in the probability of claiming a child. Liebman (2000) also examined ineligible EITC recipients, finding that between 11 and 13 percent of EITC recipients in tax year 1990 did not have a child in their household, according to the Current Population Survey. However, the author found that a large proportion of erroneous EITC payments were made to households who did have children, and that many of these households were similar to eligible households. Indications are that the rules surrounding the EITC lead to confusion regarding eligibility, with families erroneously applying for the credit due to attributes that make them close to eligible (Blumenthal, Erard, & Ho, 2005).

As a policy, the EITC provides opportunities for examining differential responses to changes in its rules that have occurred from time to time. There is a long research tradition of using these exogenous changes in the analysis of labor force response (Eissa & Liebman, 1996; Meyer & Rosenbaum, 2001), marriage rates (Rosenbaum , 2000), and fertility (Baughman & Dickert-Conlin, 2009). The latest major change to benefit levels vis-à-vis the number of children in a household occurred in 2009, when a higher benefit schedule was instituted for families with three or more children. We use this change in the EITC rules to examine the effect of EITC rules on sorting.

3. THEORETICAL AND EMPIRICAL MODEL

McCubbin (2000) presents a theoretical model for claiming fictional children that can be adapted to the choice of sorting in the face of ambiguous rules. The model begins with the choice of whether or not a filer reports a child when the filer does not actually support a child. In our case, the filer's choice is whether or not to claim a child in the household when the household also includes a filer who is more directly related to the child. McCubbin models the taxpayer's utility decision as:

$$Max(x_r, y_r): [1-p]U\big(y_t - \tau[y_r - \delta x_r] + E(y_r, x_r)\big)$$

$$+pU(y_t - \tau[y_r - \delta x_r] + E(y_r, x_r) - \pi(y_t - y_r, x_r - x_t, T, \gamma))$$

subject to $0 \leq x_r \leq 2$. In our case, the restriction on x_r is either 3, in the case of the EITC alone, or irrelevant if there is a benefit to claiming more children for a given taxpayer. In the model, y_r is reported income and x_r reported children, while the same variables with the subscript t are the true values for these measures. The probability of detection is p, τ is the tax rate, and δ is the amount of income exempt from tax for each reported child. $E(y_r, x_r)$ is the credit, π is the penalty for non-compliance, and T the tax underreport plus the overreport of EITC. The final term, γ, are demographic characteristics of the filer.

After taking the first order condition, McCubbin derives the main driver of claiming fictitious children: $\frac{\partial E(y_r, x_r)}{\partial x_r}$, which expresses the increase in tax return due to claiming a child (in the case of McCubbin, erroneously, and in our case, from optimally sorting). McCubbin points out that the penalty function in the case of evasion is crucial to the predictions of the model. In our case, we are making an assumption that the sorting behavior we are modeling is permitted by the EITC rules in most cases. However, we can assume that there is a cost to finding out about sorting, in which case a derivation of the explanatory term would be similar.

The largest return to household sorting is from the EITC and the head of household claiming status. The dependent exemptions and the Child Tax Credit are per child, while the EITC has a different schedule for additional children, and the head of household status provides a larger deduction and wider tax brackets. The return will clearly be larger for larger EITC amounts. In modeling the problem at hand, using either total household EITC amount or actual EITC receipt is problematic, since sorting/non-sorting is codetermined with either value. Therefore, we use a simulation to generate the maximum total EITC achievable for a family under all sorting possibilities. Following the model outlined above, we use the difference between this variable and the total household EITC generated through original modeling – in which the children reported in the CPS are assumed to be the children claimed. To account for heterogeneous households, we also examine in separate models those families in which at least one filer is EITC-eligible under original modeling.[4] To give a complete picture of sorting behavior, we also examine the relationship between sorting and 1) the eligibility of the reference person, 2) the eligibility of the first relative filer[5] in the household, and 3) the number of eligible filers in the household as originally modeled and via optimal sorting.

To empirically model sorting behavior, we use probit models with the explanatory variables explained above, plus reference person, state-level, and household characteristics. Reference person characteristics include adjusted gross income (AGI), filing status, age, sex, race, Hispanic origin, and education. State characteristics include the value of state EITC to the reference person under original modeling and the state minimum wage (both logged). Household characteristics include total AGI, the relationship category of the first relative —child, parent,

[4] This allows us to partially distinguish between households where doubling up occurs for economic reasons and higher income households whose adult children have not "fledged" yet. An example of the latter would be a recent college graduate with high-income parents and younger siblings who briefly returns home.
[5] There may be multiple related individual filing within a household. Relative to the reference person, the first relative is the first parent, child, or sibling of the reference person who also files a tax return.

etc.—and the age of the first relative. To control for unobservable geographic differences in tax policy, we include region fixed effects[6], and we include year fixed effects to control for any changes that occurred in tax policy over time. We cluster the standard errors on region to account for the possibility of spatial autocorrelation (Bertrand , Duflo, & Mullainathan, 2002).

Finally, using a difference-in-differences specification, we also exploit the change in EITC policy that took place in 2009. In that year, the EITC schedule for families with three or more children was increased, with a higher phase-in percentage and higher maximum credit. To our knowledge, this was the only change involving the number of dependents made in the tax rules over the period studied. We thus examine two groups before and after the rule change: those with an EITC-eligible filer under original modeling and those without. This model is identified based on the fact that the rule change would only impact sorting incentive for the eligible group. As long as the trend in sorting incentive did not change for either group except through this rule change, the model should capture its effect. In this model, the dependent variable is a 1 when a household both sorts and at least one filer in the household claims exactly three children. Because it is only possible to sort to three children if there are three or more in the household, we limit the sample to such households for this analysis.

4. DATA

The data we use for this project is the matched CPS ASEC-IRS linked files for 2005 to 2010. IRS data include the universe of 1040 filers. Census data include the CPS ASEC from 2006 to 2012, which provide information for the tax year preceding the survey year.[7]

[6] There are not enough year-state observations in the "sorter" category to include state fixed effects.
[7] Data are subject to error arising from a variety of sources. For more information on sampling and nonsampling error, see www.census.gov/prod/techdoc/cps/cpsmar12.pdf.

Records were linked in the Center for Administrative Records and Research (CARRA) at the U.S. Census Bureau. The linking process involves assigning to individuals in each data set a unique person identifier, called a Protected Identification Key (PIK). CARRA assigned these unique identifiers via the Person Identification Validation System (PVS), which employs probabilistic record linkage techniques (see Wagner and Layne (2014) for more information). CARRA uses personally identifiable information (PII) such as Social Security Number, name, date of birth, and address to assign a PIK by comparing the same fields in a master reference file constructed from federal administrative data sources. CARRA then removes the PII from the data file to anonymize the data and preserve confidentiality so it can be used for statistical purposes and research. Only those observations that received the unique key are used in the analysis.

Only certain variables from the 1040 are available for use. These include filing status, number of exemptions, wage and salary income, AGI, and number of dependent children claimed. We also have flags for certain schedules that the filer submitted, but we do not know values reported on the schedules. Importantly, we do not have a value for total tax burden or for other credits besides the EITC.

The steps in modeling EITC eligibility involve first modeling eligibility based solely on survey data. We determine the number of qualifying children that a reference person claims according to survey answers, and assume this is the number that would be claimed for tax purposes. This number is further refined by information from tax data. For example, we adjust the children claimable for a filer if the filer was claimed as a dependent by someone else, shifting the child to the person that claimed them. This is separate from sorting, as the relative filer is a dependent.

Once all of the modifications to modeled children are made[8], we restrict the sample to only those households where there was at least one other adult tax filer who was related to, but not married to, the reference person. If the reference person was married and the spouse filed, all of the tax information was applied to the reference person. We then appended information from each other related filer in the household to the reference person (now also called the reference filer). In the vast majority of cases, there is a single adult relative in the household, whom we call the "first relative filer." For our created variable, "sorter," we calculated the difference between the modeled children and the 1040 claimed children for the reference filer, and the difference between modeled children and claimed children for each related filer. If the difference was less than 0 for the reference filer and greater than 0 for a related filer, we considered that household to have sorted children. Furthermore, if the difference was greater than 0 for the reference filer and less than 0 for a related filer, we considered that household to have sorted. In all analyses, the reference person in each household is the unit of observation, and our dependent variable is equal to 1 for reference filers in households that sorted.

The main independent variable of interest is the optimal tax burden a household can achieve through sorting children. While we do not have a full tax model that can be run on restricted data at this time, the simulations on artificial data using TAXSIM showed that the optimal household tax burden is strongly determined by EITC amount. We use the eligibility modeling described before to create all of the possible EITC credit outcomes for each filer in the household if children are sorted. In the data, we observe sorting between the first filer and a maximum of three other relative filers, so we limit the number of modeling repetitions to 4 filers and 6 children. The number of times a simulation is run for a household is based on the possible

[8] For a full description of how our EITC eligibility modeling unfolds, see Plueger (2009).

combinations of filers and children, which is $\frac{(n+r-1)!}{n!(r-1)!}$, where n is the number of children and r is the number of filers.[9] For each combination, we capture the total household EITC achieved through that sorting, and in the end save the maximum amount possible. The difference between this and the total original modeled amount for the household is the main explanatory variable under consideration.

5. RESULTS

5.1 Summary statistics

We first looked at the predictors of whether or not a household included multiple related adult tax filers, and whether the incidence of multiple-related-filer households increased between tax years 2005 and 2010. We limited the sample to households in which at least one dependent child was modeled for someone in the household. Table 1 shows means for the variables used in the analysis and presents t statistics indicating whether or not the means are different between households with and without multiple adult tax filers. Reference persons in homes with multiple related filers tend to have lower AGI and to be older, and are more likely to be single, female, Black alone, Asian alone, and other or mixed race . They are less likely to be White alone or married. Finally, they tend to have lower levels of education than reference persons not in multiple-related-filer households. Panel B shows that the rate of multiple-related-filer households increased between 2005 and 2010 (changing from 10.5 percent of tax-filing households with one child or more to 12.6 percent), consistent with studies on the subject.

[9] The maximum number of repetitions is thus 84.

Table 1. Summary statistics of variables used in model, households with multiple filers and those without

Panel A	Without a relative filer	With a relative filer	t-statistic
AGI of reference person (log)	10.89	10.51	39.37
Dependent children	1.94	1.84	13.09
Single	0.04	0.19	-81.96
Married	0.72	0.59	38.09
Head of household	0.23	0.22	4.86
Age of reference person	40.60	48.28	-100.56
Female reference person	0.50	0.53	-7.77
White alone	0.83	0.76	24.37
Black alone	0.10	0.14	-15.65
Asian alone	0.04	0.06	-14.08
Other race	0.03	0.04	-9.69
Hispanic	0.13	0.21	-29.27
Less than high school	0.09	0.16	-33.29
High school graduate	0.26	0.33	-20.81
Some college	0.31	0.30	0.75
BA/BS or more	0.35	0.21	39.69

Panel B	Without a relative filer	With a relative filer
2005	89.47	10.53
2006	89.08	10.92
2007	87.28	12.72
2008	87.41	12.59
2009	87.90	12.10
2010	87.42	12.59
Total	88.10	11.90
Observations	147,229	19,897

Source: CPS ASEC-IRS linked file for tax years 2005–2010. The unit of observation is the CPS reference person. Included are all reference persons who filed a 1040 and who had a dependent child in their household. The means of the variables listed in the first column are reported in Panel A. Panel B reports the percentage of households in each year in each category.

In Table 2, multiple-filer households with reference persons who sort children are compared with those who do not. The incidence of sorting over the population of multifamily households is about 11 percent. Reference persons in sorting households are more likely than non-sorting reference persons to be single or head of household, and they are less likely to married. In terms of race, sorters are more likely than non-sorters to be Black alone or other race, and they are less likely to be White or Asian alone. Sorters also have lower AGI and are older. In

terms of educational attainment, sorters are more likely than non-sorters to have a high school education or less, while they are less likely than non-sorters to have a college degree. In terms of household characteristics, sorting households have more filers and more children than do non-sorting households, and they are less likely to have a child as the first filing relative and more likely to have a grandchild, sibling, or other relative.

Table 2. Summary statistics, sorters versus non-sorters among multi-filer households

	Non-sorter	Sorter	t-statistic
AGI of reference person, log	10.56	10.12	11.51
Single	0.19	0.23	-4.49
Married	0.61	0.45	14.06
Head of household	0.21	0.32	-12.44
Age of reference person	48.17	49.14	-3.26
Female reference person	0.52	0.58	-5.60
White alone	0.77	0.67	10.49
Black alone	0.12	0.23	-13.48
Asian alone	0.06	0.05	3.23
Other race	0.04	0.06	-3.23
Hispanic	0.20	0.28	-8.86
Less than HS	0.15	0.24	-10.92
High school graduate	0.33	0.37	-3.38
Some college	0.31	0.29	1.92
BA/BS or more	0.22	0.11	11.61
Total AGI (log)	9.65	9.67	-0.68
Child	0.70	0.66	3.32
Grandchild	0.01	0.03	-5.58
Parent	0.13	0.12	1.34
Sibling	0.06	0.08	-2.45
Other	0.10	0.12	-2.27
Age of first relative filer	31.21	31.00	0.61
Number of dependent children	1.76	2.45	-28.28
Number of filers	1.17	1.25	-8.11
Observations	17,729	2,150	

Source: CPS ASEC-IRS linked file for tax years 2005–2010. The unit of observation is the CPS reference person. Included are all reference persons who filed a 1040, who had a relative filer in their household, and who had a dependent child in their household.

This brings us to the make-up, in terms of relatives, of sorting and non-sorting households. Table 3 shows the type of relative reported by the reference person to the CPS for

the "first relative" and "second relative" in multiple-related-filer households. The only reason why some relatives are labeled "first" and others are "second" is due to the ordering of the coding for the variable in the survey data. For the vast majority of both groups, only one other adult related filer lives in the household. Non-sorting households were more likely to have an adult child or a parent in the household, while sorters were more likely to have a grandchild or a relative in the "other" category. Sorters have a higher rate of having a second adult related filer compared with non-sorters (21 percent versus 14 percent). Households with two filing children represented about 8 percent of non-sorting households and 12 percent of sorting households. Other configurations that were represented more than 1 percent of the time were households with a parent and a sibling filer, and those with a child and a relative in the "other" category. All other configurations were represented less than 1 percent of the time in either sorting or non-sorting households. Sorters were more likely to have two or more relatives in the household, but the rates for certain configurations were statistically different between sorters and non-sorters while others were not. The rates that differ were households with two children, a child and a relative from the "other" category, a child and a grandchild, two siblings,

Table 3. Relative composition of sorting versus non-sorting households

	Non-sorter	Sorter
Child alone	58.90	50.19
Grandchild alone	1.09	2.84
Parent alone	10.45	8.65
Sibling alone	5.48	5.91
Other alone	9.67	11.16
Subtotal	85.59	78.75
Child-child	8.22	12.00
Parent-sibling	1.59	1.91
Child-other	1.37	2.09
Other-other	0.52	0.60
Child-grandchild	0.50	0.98
Parent-other	0.40	0.70
Child-sibling	0.34	0.37
Sibling-sibling	0.33	0.70
Sibling-other	0.32	0.88
Child-parent	0.30	0.51
Parent-parent	0.24	0.42
Subtotal	14.40	21.25
Observations	17,729	2,150

Source: CPS ASEC-IRS linked file for tax years 2005–2010. The unit of observation is the CPS reference person. Included are all reference persons who filed a 1040, who had a relative filer in their household, and who had a dependent child in their household. Relative filers were listed in order in the data based on their CPS code (thus, "child" was coded "4," "parent" coded "5," and so on). Categories were defined based on combinations of the first relative filer and the second relative filer listed. Categories that do not appear were represented by fewer than 6 persons in the "sorter" category; thus, the subtotal percentages do not sum to 100 percent.

or one sibling and a relative from the other category. Other configurations did not differ between types of households.

5.2 Probit results

The marginal effects from the probit models described in Section 3 are shown in Table 4. In each model, the dependent variable is 1 for reference persons in households that sort. The main explanatory variables are: modeled eligibility for the reference person or any related adult filer based on our original EITC eligibility modeling (models 1-4); the maximum number of related filers in a household possible under optimal sorting (models 5 and 6); the maximum total household EITC possible under optimal sorting (models 7 and 8); the difference between total household EITC and maximum possible EITC under optimal sorting (models 9 through 12). Odd-numbered models include characteristics for the reference person only, while even-numbered models include characteristics for the first relative filer and the household. The sample for models 11 and 12 is restricted to those households that included an EITC-eligible filer under original modeling.

Looking at the results of models 1 through 4, we see that a change for the reference person from non-eligible for EITC to eligible (under original modeling) is associated with a 0.04 increase in the probability of sorting when only reference person characteristics are included. This decreases slightly, to 0.03, when household characteristics are added to the model.

Table 4. Probit models predicting sorting behavior. Dependent variable equals 1 when a household sorts and 0 otherwise

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Eligible for EITC, reference person	0.05***	0.03***										
	(0.00)	(0.01)										
Eligible for EITC, relative			0.05***	0.04***								
			(0.00)	(0.01)								
Maximum filers eligible					0.05***	0.03***						
					(0.00)	(0.00)						
Max per person EITC (log)							0.01***	0.01***				
							(0.00)	(0.00)				
EITC difference per person (log)									0.00	0.00	0.01*	0.01*
									(0.00)	(0.00)	(0.00)	(0.00)
Reference person characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Household characteristics	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
Observations	19,878	19,878	19,878	19,878	19,878	19,878	19,878	19,878	19,878	19,878	9,017	9,017

* $p<0.05$, ** $p<0.01$), *** $p<0.001$. Source: CPS ASEC--IRS linked file for tax years 2005--2010. Marginal effects are reported; clustered standard errors in parentheses. Also included in even-numbered models were fixed effects for year and region. The unit of analysis is the CPS reference person, and included are 1040-filing reference persons with at least one other adult related filer and at least one dependent child in the household. Columns 11 and 12 include only households in which a filer was modeled as eligible for the EITC. Eligibility for EITC is determined using the survey-reported number of dependents for the filers in the household. The dependent variable marks households that sort children among adult related 1040 tax filers, according to the definition described in the text.

When any other member of the household is eligible, the likelihood of sorting increases by about 0.05 when only reference filer characteristics are included and 0.04 when household characteristics are included. Turning to the models using dependent variables generated by our simulation, model 5 shows that as the number of potentially eligible filers in the household increases by 1, the propensity to sort increases by about 0.05 (0.03 when household characteristics are included). As predicted, the larger the possible household EITC, the more likely it is that the household will sort, and this result holds true for all households. Models 7 and 8 show that a 10 percent increase in simulated household EITC is associated with a 0.01 greater likelihood that a household will sort.

Interestingly, models 9 and 10 show that the *difference* between modeled and optimal EITC is not associated with sorting when looking at all households, with a reported marginal effect and standard error of 0. It should be noted, however, that we include all households with adult related filers, including those in which no filer is originally modeled as eligible for the EITC. When we limit the analysis to only those households where at least one filer was eligible for EITC under original modeling, the coefficient on the difference is positive and statistically significant, indicating that a 10 percent increase in the difference is associated with about a 0.01 increase in sorting. These results, when taken together, may indicate the influence of information. In other words, sorting may occur only in households where at least one filer determines on his own—or through a paid preparer—that he is eligible for EITC, and the sorting takes place after the information is generated. The sample restriction may separate out cases in which the reference filers are not eligible for EITC due to high adjusted gross income, but have adult, filing children and younger children living in the same home. In such a household, if the

reference filer is never eligible for EITC due to income, and a relative filer is never eligible due to not having children, the household may not have the information necessary to consider sorting. [10]

Table 5 provides some evidence that the economic relationship between filers in sorting and non-sorting households may be different. For reference filers and first relative filers in sorting households, earnings are considerably less than in non-sorting households. Non-sorting reference filers on average earn $22,000 more than sorting reference filers. More important, the difference between the earnings of reference filers and first relative filers in non-sorting

Table 5. Earnings and differences in earnings between sorters and non-sorters

	Non-sorters	Sorters	t-statistic
Reference filer earnings	55,057.23	33,742.74	10.67
First relative filer earnings	20,351.31	18,321.55	3.14
Second relative filer earnings	20,138.69	19,775.68	0.35
Reference minus first filer earnings	36,125.81	17,127.17	8.96
Observations	17,729	2,150	

Source: CPS ASEC-IRS linked file for tax years 2005–2010. The unit of observation is the CPS reference person. Included are all reference persons who filed a 1040, who had a relative filer in their household, and who had a dependent child in their household. Earnings were calculated using the 1040 wage and salary earnings, supplemented with W-2 wage information.

households is much larger than in sorting households, indicating that reference filers in non-sorting households may see themselves as financial providers to these relatives, even if they no longer claim them for tax purposes.

5.3 Sorting to exactly three

Table 6 shows the results of the difference-in-differences analysis. As in Table 4, the odd-numbered models include only reference-person characteristics while the even-numbered include household characteristics. Here, the dependent variable is a 1 when a household both sorts and at least one filer in the household claims exactly three children. The unit of analysis is again the reference person, and the standard errors are corrected by clustering at the region level.

[10] The full set of results is reported in appendix Table A1. Also provided are the same models as the probits, but using OLS, in Table A2. The results between the probit models and linear probability models are indistinguishable.

Because it is not possible to claim three children unless there are three or more children in the household, we limit the sample to households that meet this definition. We also control for the number of children—three or more—in the remaining households. All other control variables from the preceding model were included.

Table 6. Difference-in-differences model; dependent variable equals 1 when a household sorts and at least one household filer claims three children, and 0 otherwise

	(1)	(2)	(3)	(4)	(5)	(6)
Eligible*post	0.07***	0.07***				
	(0.01)	(0.01)				
Post	0.02	0.01				
	(0.01)	(0.01)				
Eligible	-0.00	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Eligible*2006			0.000	0.00	0.00	0.00
			(0.02)	(0.02)	(0.02)	(0.02)
Eligible*2007			0.01	0.01	0.01	0.01
			(0.01)	(0.01)	(0.01)	(0.01)
Eligible*2008			-0.00	-0.00	-0.00	-0.00
			(0.01)	(0.01)	(0.01)	(0.01)
Eligible*2009			0.08*	0.08*	0.08*	0.08**
			(0.023)	(0.021)	(0.023)	(0.021)
Eligible*2010			0.06*	0.07*	0.06*	0.07*
			(0.022)	(0.022)	(0.022)	(0.022)
Reference person characteristics	yes	yes	Yes	yes	yes	yes
Household characteristics	no	yes	No	yes	no	yes
Year fixed effect	no	no	Yes	yes	yes	yes
Linear time trend	no	no	No	no	yes	yes
Observations	4,039					

* *p*<0.05, ** *p*<0.01, *** *p*<0.001. Source: CPS ASEC--IRS linked file for tax years 2005--2010. Clustered standard errors in parentheses. Also included were region fixed effects. The unit of analysis is the CPS reference person, and included are 1040-filing reference persons with at least one other adult related filer and at least three dependent children in the household. The dependent variable is an interaction of ``sorter,'' defined in the text, and a marker equal to 1 when any filer in the household claims exactly three children. The first three rows describe the results of a difference-in-differences model where the post period is defined as year 2009 or later.

In the simple difference-in-differences, reported in columns 1 and 2, the coefficient of interest is the interaction term *Eligible*post*, which indicates how much sorting to exactly three children increased in the post period for EITC eligible households compared with EITC-ineligible households. In the post period, sorting to three children increased about 7 percent for eligible households compared to those ineligible.[11]

[11] The full results appear in appendix Table A3.

In the interest of completeness, some alternative specifications are offered. Columns 3 and 4 report results when the eligibility marker is interacted with a year fixed effect, giving the difference in reporting by year. The coefficients on the interaction terms indicate that the increase in sorting to exactly three did not differ in a significant way between 2009 and 2010, the two years in the data when the rule was in effect. The coefficients from models 2 and 4 do not change in models 5 and 6, when we include a linear time trend in the model.

This increase in sorting to three children provides some supporting evidence that sorting in general is a behavior that occurs in direct response to tax rules, as opposed to being an artifact of the optimization exercise. Because the sorting incentive changed only for EITC-eligible families, the most plausible explanation for the increase is that tax filers—or, more likely, tax preparers—took advantage of the new rule to improve the return for large households.

6. CONCLUSION

The analysis presented in this paper is the first that we know of to examine the sorting of dependents in multi-family households. The results add to our knowledge about a particular type of tax-avoidance behavior—that of intra-family arbitrage—that has in general received little attention in the literature. This work also contributes to research recently conducted on tax-filing behavior that takes advantage of discontinuities and other incentives in the income tax laws of the United States. Our topic is particularly relevant at present due to the increasing incidence of households "doubling up" during the Great Recession.

Our results indicate that households with more potential EITC-eligible filers and more children are also more likely to sort. Moreover, households with at least one filer who was eligible for the EITC in initial modeling were more likely to sort when the return, measured as

maximum possible EITC minus total modeled EITC, was larger. However, in households where no one was eligible for EITC in initial modeling, no sorting occurred in connection with an EITC increase. This gives some indication of an information effect, since presumably higher income filers or those without children are not likely to know about eligibility rules. In looking at the difference in earnings between filers in households, sorting was associated with households where filers' earnings were closer in value, indicating that the economic relationship between filers in sorting households differs from those in non-sorting households.

Our analysis has several limitations. First, based on the survey data available, we are making an assumption that all households with multiple adult filers live together for the purposes of combining resources, and that they lived together for the tax year. Thus, we are comparing households that are truly multi-family over the tax year with those that may simply have a relative staying with them for a short time without any combining of household resources. It is difficult to distinguish between two such households without panel data, since both the survey data and the tax data provide information on where a filer was living at or around the time of filing Moreover, our analysis begins and ends with our eligibility modeling for the EITC. Using the information given to us by survey respondents, we do our best to correctly assign dependents, and then compare that information to what the respondents claimed on their taxes. While we take the survey information as the true value in regards to which dependents belong to whom, there is the chance that the tax information reported on the 1040 is actually the true value. On the other hand, an analysis of sorting behavior when exactly three children are claimed provides some support that we are finding a true effect and not an artifact of our data or our modeling.

In spite of these limitations, the results of the analysis add to our understanding of how multi-family households navigate the tax and transfer system and use quirks in the rules to their

advantage. Because the returns to sorting are often large, differences between refunds under sorting may mean that a household escapes the standard definition of poverty. The fact that sorting occurs only in households where at least one filer is originally modeled as eligible for EITC further adds to our understanding of how tax rules are disseminated within households, and between tax preparers and filers. And last, as with any tax avoidance, the avoidance of income taxes through dependent sorting has implications for public finance and equity in taxation. Understanding this behavior should inform any investigation into the resources of complex households, including modeling for the alternative poverty measure or determining the nature of responsibility for children in households that pool resources.

Works Cited

Andreoni, J., Erard, B., & Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature*, 818–860.

Baughman, R., & Dickert-Conlin, S. (2009). The Earned Income Tax Credit and Fertility. *Journal of Population Economics, 22*(3), 537-563.

Bertrand , M., Duflo, E., & Mullainathan, S. (2002). *How much should we trust differences-in-differences estimates?* National Bureau of Economic Research.

Blumenthal, M., Erard, B., & Ho, C.-C. (2005). Participation in and compliance with the Earned Income Tax Credit. *National Tax Journal*, 189-213.

Chetty, R., & Saez, E. (2009). *Teaching the tax code: Earnings responses to an experiment with EITC recipients.* Tech. rep., National Bureau of Economic Research.

Chetty, R., Friedman, J. N., & Saez, E. (2012). *Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings.* National Bureau of Economic Research.

Duflo, E. (2003). Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa. *The World Bank Economic Review*, 1–25.

Eissa, N., & Liebman, J. (1996, May). Labor Supply Responses to the Earned Income Tax Credit. *Quarterly Journal of Economics*, 605-637.

Ellwood, D. T., & Liebman, J. B. (2001). The middle-class parent penalty: Child benefits in the U.S. tax code. In *Tax Policy and the Economy* (Vol. 15, pp. 1-40). MIT Press.

Feenberg, D., & Coutts, E. (1993). An Introduction to the TAXSIM model. *Journal of Policy Analysis and Management, 12*(1), 189-194.

Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics*, 674–680.

Haurin, D. R., Hendershott, P. H., & Kim, D. (1993). The impact of real rents and wages on household formation. *The Review of Economics and statistics*, 284-293.

Jones, M. R. (2013). Changes in EITC Eligibility and Participation, 2005-2009. *IRS Research Bulletin*.

Jones, M. R. (2013). *The EITC and Labor Supply: Evidence from a Regression Kink Design.* U.S. Census Bureau.

Kopczuk, W., & Pop-Eleches, C. (2007). Electronic filing, tax preparers and participation in the Earned Income Tax Credit. *Journal of Public Economics, 91*(7), 1351–1367.

LaLumia, S., & Sallee, J. M. (2012). The value of honesty: empirical estimates from the case of the missing children. *International Tax and Public Finance, 20*(2), 192-224.

Liebman, J. B. (2000). Who are the ineligible EITC recipients. *National Tax Journal, 53*(4), 1165-1185.

McCubbin, J. (2000). EITC noncompliance: The determinants of misreporting children. *National Tax Journal, 53*(4), 1135-1164.

Meyer, B., & Rosenbaum, D. (2001). Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers. *Quarterly Journal of Economics*.

Mykyta, L., & Maccartney, S. (2011). *The effects of recession on household composition: 'Doubling up' and economic well-being.* US Census Bureau.

Ohlsson, H. (2007). *Tax avoidance—a natural experiment.* Working Paper, Department of Economics, Uppsala University.

Plueger, Dean. (2009) *Earned Income Tax Credit Participation Rate for Tax Year 2005.* IRS Research Bulletin

Rosenbaum , D. (2000). *Taxes, the Earned Income Tax Credit, and Marital Status.* Joint Center for Poverty Research Working Paper (177).

Saez, E. (2010). Do Tax Filers Bunch at Kink Points? Evidence, Elasticity, and Salience Effects. *American Economic Journal: Economic Policy, 2*(3), 180-212.

Scholz, J. K. (1994). The earned income tax credit: Participation, compliance, and antipoverty effectiveness. *National tax journal*, 63–87.

Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *The Journal of Economic Perspectives*, 25–48.

Slemrod, J., & Yitzhaki, S. (2002). Tax avoidance, evasion, and administration. In *Handbook of Public Economics 3* (pp. 1423-1470).

Stephens Jr, M., & Ward-Batts, J. (2004). The impact of separate taxation on the intra-household allocation of assets: evidence from the UK. *Journal of Public Economics*, 1989–2007.

Stiglitz, J. E. (1988). *The general theory of tax avoidance.* National Bureau of Economic Research Cambridge.

Wagner, D., & Layne, M. (2012). *Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' Record Linkage Software.* Washington, DC: Center for Administrative Records Research and Applications Internal Document, US Census Bureau.

Appendix: Expanded material

Table A1: Full probit models predicting sorting behavior. Dependent variable equals 1 when a household sorts and 0 otherwise

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Eligible for EITC, reference person	0.05***	0.03***										
	(0.00)	(0.01)										
Eligible for EITC, relative			0.05***	0.04***								
			(0.00)	(0.01)								
Maximum filers eligible					0.05***	0.03***						
					(0.00)	(0.00)						
Max per person EITC (log)							0.01***	0.01***				
							(0.00)	(0.00)				
EITC difference per person (log)									0.00	0.00	0.01*	0.01*
									(0.00)	(0.00)	(0.00)	(0.00)
AGI of reference person	0.00	0.00	0.00	0.00	0.00	0.00	0.00*	0.00	0.00*	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Single	0.03***	0.04***	0.03***	0.04***	0.03***	0.04***	0.03***	0.03***	0.03***	0.03***	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Head of Household	0.04***	0.04***	0.05***	0.05***	0.04***	0.04***	0.05***	0.05***	0.05***	0.05***	0.04***	0.04***
	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age of reference person	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female reference person	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Black alone	0.05***	0.05***	0.05***	0.05***	0.05***	0.05***	0.05***	0.05***	0.06***	0.06***	0.05***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Asian alone	0.00	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Other race	0.03***	0.02*	0.03***	0.02	0.03***	0.02*	0.04***	0.03**	0.04***	0.03**	0.03	0.03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Hispanic	0.04***	0.02**	0.03***	0.02*	0.03***	0.02*	0.04***	0.03**	0.04***	0.03***	0.01	0.01
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
High School graduate	0.03***	0.02*	0.03***	0.02**	0.03***	0.02*	0.03***	0.03***	0.03***	0.03***	0.03**	0.03*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Some college	0.04***	0.03***	0.05***	0.03***	0.04***	0.03***	0.05***	0.04***	0.05***	0.04***	0.05***	0.04***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
BA/BS or more	0.08***	0.07***	0.09***	0.07***	0.08***	0.07***	0.09***	0.08***	0.09***	0.09***	0.09***	0.08***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Total AGI (log)		0.00		0.00		0.00		0.00		0.00*		0.00
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Child		0.02		0.02		0.02*		0.03*		0.03		0.01
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.02)
Grandchild		0.01		0.01		0.01		0.00		0.00		0.00
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.05)
Parent		0.00		0.00		0.00		0.01		0.01		0.01
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.03)

Other		0.01		0.01		0.01		0.02		0.01		0.03
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)
Age of first relative		0.00		0.00		0.00		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
State EITC (log)		0.00		0.00***		0.00		0.01***		0.01***		0.00
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
State minimum wage (log)		0.04		0.04		0.04		0.04		0.04		0.09*
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.05)
Year=2006		0.02		0.02*		0.02		0.02*		0.02*		0.04*
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.02)
Year=2007		0.02		0.02		0.02		0.02		0.02		0.04
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.02)
Year=2008		0.03*		0.03*		0.03*		0.03*		0.03*		0.06**
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.02)
Year=2009		0.03*		0.03*		0.03*		0.03*		0.03		0.06**
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.02)
Year=2010		0.03*		0.03		0.03		0.03*		0.03		0.06***
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)
New England		0.02***		0.02***		0.02***		0.02***		0.02***		0.04***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Southeast		0.02***		0.02***		0.02***		0.02***		0.03***		0.01*
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
East Central		0.01***		0.02***		0.01***		0.02***		0.02***		0.01**
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
North Central		0.01***		0.01***		0.01***		0.01**		0.00		0.01*
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Southwest		0.03***		0.03***		0.03***		0.04***		0.04***		0.02***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.01)
West		0.00		0.01		0.00		0.02***		0.02***		0.00
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.01)
Number of dependent children		0.04***		0.04***		0.04***						
		(0.00)		(0.00)		(0.00)						
Number of filers		0.02***		0.02***		0.02***						
		(0.00)		(0.00)		(0.00)						
Observations	19,878	19,878	19,878	19,878	19,878	19,878	19,878	19,878	19,878	19,878	9,016	9,016

* p < 0:05, ** p < 0:01, *** p < 0:001. Source: CPS ASEC-IRS linked file for tax years 2005-2010. Marginal effects are reported; clustered standard errors in parentheses. The unit of analysis is the CPS reference person, and included are 1040-filing reference persons with at least one other adult related filer and at least one dependent child in the household. Columns 11 and 12 include only households in which a filer was modeled as eligible for the EITC. Eligibility for EITC is determined using the survey-reported number of dependents for the filers in the household. The dependent variable marks households that sort children among adult related 1040 tax filers, according to the definition described in the text.

Table A2: OLS models predicting sorting behavior. Dependent variable equals 1 when a household sorts and 0 otherwise

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Eligible for EITC, reference person	0.05***	0.04***										
	(0.01)	(0.01)										
Eligible for EITC, relative			0.06***	0.04***								
			(0.00)	(0.01)								
Maximum filers eligible					0.05***	0.03***						
					(0.00)	(0.00)						
Max per person EITC (log)							0.01**	0.01**				
							(0.00)	(0.00)				
EITC difference per person (log)									0.00	0.00	0.01*	0.01*
									(0.00)	(0.00)	(0.00)	(0.00)
AGI of reference person	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Single	0.03**	0.04***	0.02*	0.04**	0.03**	0.04**	0.03**	0.02*	0.03**	0.02*	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Head of Household	0.04***	0.05***	0.05***	0.05***	0.05***	0.05***	0.05***	0.05***	0.06***	0.05***	0.04**	0.04*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age of reference person	0.00**	0.00***	0.00**	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female reference person	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Black alone	0.07**	0.06**	0.06**	0.06**	0.06**	0.06**	0.07**	0.06**	0.07**	0.07**	0.05*	0.05*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Asian alone	0.00	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Other race	0.04*	0.02	0.04*	0.02	0.04*	0.02	0.04*	0.03*	0.04*	0.03*	0.03	0.03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Hispanic	0.04***	0.02*	0.04***	0.02*	0.04***	0.02*	0.04***	0.03**	0.04***	0.03**	0.01	0.01
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
High School graduate	0.03**	0.02*	0.04**	0.02*	0.03**	0.02*	0.04**	0.04**	0.04**	0.04**	0.03*	0.03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Some college	0.05***	0.03**	0.05***	0.04**	0.05***	0.03**	0.05***	0.05***	0.05***	0.05***	0.05*	0.05*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
BA/BS or more	0.08***	0.07***	0.08***	0.07***	0.07***	0.07***	0.08***	0.08***	0.09***	0.08***	0.08**	0.08**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Total AGI (log)		0.00		0.00		0.00		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Child		0.02		0.02		0.03		0.03		0.03		0.01
		(0.02)		(0.02)		(0.01)		(0.02)		(0.02)		(0.02)
Grandchild		0.02		0.02		0.02		0.01		0.02		0.00
		(0.03)		(0.03)		(0.03)		(0.03)		(0.03)		(0.06)
Parent		0.00		0.00		0.01		0.01		0.01		0.01
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.03)
Other		0.01		0.01		0.01		0.02		0.02		0.03
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.03)

Age of first relative		0.00		0.00		0.00		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
State EITC (log)		0.00		0.00*		0.00		0.01**		0.01**		0.00
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
State minimum wage (log)		0.04		0.04		0.04		0.04		0.04		0.09
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.05)
Year=2006		0.02		0.02		0.02		0.02		0.02		0.04
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.02)
Year=2007		0.02		0.02		0.02		0.02		0.02		0.03
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.02)
Year=2008		0.03		0.03		0.03		0.03*		0.03		0.05*
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.02)
Year=2009		0.03		0.03		0.03		0.03		0.02		0.05*
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.02)
Year=2010		0.03		0.03		0.03		0.03		0.03		0.06**
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)
New England		0.02**		0.02**		0.02**		0.02**		0.02**		0.04***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Southeast		0.02***		0.03***		0.02***		0.03***		0.03***		0.01*
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
East Central		0.01***		0.02***		0.01***		0.02***		0.02***		0.01
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
North Central		0.01*		0.01*		0.01*		0.00		0.00		0.01
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.01)
Southwest		0.03***		0.03***		0.03***		0.04***		0.04***		0.03**
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.01)
West		0.00		0.01		0.00		0.01*		0.02*		0.00
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.01)
Number of dependent children		0.06***		0.06***		0.05***						
		(0.00)		(0.00)		(0.00)						
Number of filers		0.02**		0.02**		0.02**						
		(0.00)		(0.00)		(0.01)						
Constant	0.08	0.15*	0.12*	0.12*	0.03	0.18**	0.09*	0.03	0.13**	0.03	0.13**	0.03
	(0.04)	(0.05)	(0.04)	(0.05)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.05)	(0.02)	(0.10)
Observations	19,878	19,878	19,878	19,878	19,878	19,878	19,878	19,878	19,878	19,878	9,016	9,016

* p < 0:05, ** p < 0:01, *** p < 0:001. Source: CPS ASEC-IRS linked file for tax years 2005-2010. Coefficients are reported; clustered standard errors in parentheses. The unit of analysis is the CPS reference person, and included are 1040-filing reference persons with at least one other adult related filer and at least one dependent child in the household. Columns 11 and 12 include only households in which a filer was modeled as eligible for the EITC. Eligibility for EITC is determined using the survey-reported number of dependents for the filers in the household. The dependent variable marks households that sort children among adult related 1040 tax filers, according to the definition described in the text.

Table A3: Difference-in-differences models; dependent variable equals 1 when a household sorts and at least one household filer claims three children, 0 otherwise

	(1)	(2)	(3)	(4)	(5)	(6)
Eligible*post	0.07**	0.07***				
	(0.01)	(0.01)				
Post	0.02	0.01				
	(0.01)	(0.01)				
Eligible	-0.00	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Eligible*2006			0.00	0.00	0.00	0.00
			(0.02)	(0.02)	(0.02)	(0.02)
Eligible*2007			0.01	0.01	0.01	0.01
			(0.01)	(0.01)	(0.01)	(0.01)
Eligible*2008			-0.00	-0.00	-0.00	-0.00
			(0.01)	(0.01)	(0.01)	(0.01)
Eligible*2009			0.08*	0.08**	0.08*	0.08**
			(0.02)	(0.02)	(0.02)	(0.02)
Eligible*2010			0.06*	0.07*	0.06*	0.07*
			(0.02)	(0.02)	(0.02)	(0.02)
AGI of reference person	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Single	0.00	0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Head of Household	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age of reference person	0.00	0.00*	0.00	0.00*	0.00	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female reference person	0.01	0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Black alone	0.01	0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Asian alone	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Other race	0.01	0.01	0.01	0.01	0.01	0.01
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Hispanic	0.02	0.02	0.03	0.03	0.03	0.03
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
High School graduate	0.00	0.01	0.00	0.01	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Some college	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
BA/BS or more	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
New England	0.01*	0.00	0.00*	-0.00	0.00*	-0.00
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
Southeast	-0.01***	-0.02**	-0.01***	-0.01*	-0.01***	-0.01*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
East Central	0.01***	0.00*	0.01***	0.01*	0.01***	0.01*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
North Central	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Southwest	0.00*	0.00	0.00*	0.01	0.00*	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
West	0.02**	0.01*	0.02**	0.01	0.02**	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Total AGI (log)		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)
Child		-0.02		-0.02		-0.02
		(0.03)		(0.03)		(0.03)
Grandchild		-0.05		-0.04		-0.04

		(0.04)		(0.04)		(0.04)
Parent		-0.01		-0.01		-0.01
		(0.03)		(0.03)		(0.03)
Other		-0.02		-0.02		-0.02
		(0.03)		(0.03)		(0.03)
Age of rel 1		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)
Number of dependent chidren		0.03***		0.03***		0.03***
		(0.00)		(0.00)		(0.00)
Number of filers		-0.01**		-0.01**		-0.01**
		(0.00)		(0.00)		(0.00)
lnsteitc		-0.00		-0.00		-0.00
		(0.00)		(0.00)		(0.00)
State minimum wage (log)		0.02		0.05		0.05
		(0.02)		(0.04)		(0.04)
Year=2006			-0.00	-0.01	-0.00	-0.00
			(0.01)	(0.01)	(0.01)	(0.01)
Year=2007			-0.01*	-0.02*	-0.02	-0.02
			(0.01)	(0.01)	(0.01)	(0.01)
Year=2008			-0.01	-0.02	-0.01	-0.01
			(0.01)	(0.01)	(0.02)	(0.02)
Year=2009			0.01	-0.01	0.00	0.01
			(0.01)	(0.02)	(0.02)	(0.02)
Year=2010			0.01	-0.01		
			(0.02)	(0.02)		
Trend					0.00	-0.00
					(0.00)	(0.00)
Constant	-0.01	-0.16*	-0.01	-0.22*	-0.01	-0.21*
	(0.01)	(0.05)	(0.01)	(0.08)	(0.01)	(0.07)
Observations	4,039	4,039	4,039	4,039	4,039	4,039