The Antipoverty Impact of the EITC: New Estimates from Survey and Administrative Tax Records

by

Maggie R. Jones
U.S. Census Bureau

James P. Ziliak
University of Kentucky

CES 19-14 April, 2019

Abstract

Evaluations of the EITC, including its antipoverty effectiveness, are based on simulated EITC benefits using either the Census Bureau's tax module or from external tax simulators such as the National Bureau of Economic Research's TAXSIM or Jon Bakija's model. Each simulator utilizes model-based assumptions on who is and who is not eligible for the EITC, and conditional on eligibility, assumes that participation is 100 percent. However, recent evidence suggests that take-up of the EITC is considerably less than 100 percent, and thus claims regarding the impact of the program on measures of poverty may be overstated. We use data from the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) linked to IRS tax data on the EITC to compare the distribution of EITC benefits from three tax simulation modules to administrative tax records. We find that significantly more actual EITC payments flow to childless tax units than predicted by the tax simulators, and to those whose family income places then well above official poverty thresholds. However, actual EITC payments appear to be target efficient at the individual tax unit level, whether correctly paid or not. We then compare the antipoverty impact of the EITC across the survey and administrative tax measures of EITC benefits. We find that in the full CPS ASEC the tax simulators overestimate the antipoverty effects of the EITC by about 1.8 million persons in a typical year. Restricting to a harmonized sample of filers, we find that the antipoverty estimates derived from the TAXSIM and Bakija models align more closely to actual EITC payments compared to the CPS, suggesting a discrepancy in assignment of tax filers between the tax simulators.

[*] Address correspondence to Maggie Jones at margaret.r.jones@census.gov. We are grateful for the assistance of Josh Mitchell on the internal Census files with TAXSIM, and to Jon Bakija for sharing his tax simulator. We benefitted from helpful comments on early versions from Bruce Meyer, Jon Rothbaum, and seminar participants at the fall 2017 Association of Public Policy and Management Research Conference, the Census Bureau, and the University of Illinois at Chicago.

The Earned Income Tax Credit (EITC) is one of the largest and most studied antipoverty programs in the United States, with extensive research on labor supply, consumption, marriage, fertility, child achievement, and poverty alleviation.[1] Spending on the EITC exceeded $68 billion in 2013 (Falk and Crandall-Hollick 2016), and the program was estimated to lift more persons out of poverty than any other safety net program for children and nonelderly working households (Hoynes and Patel 2017; Ziliak 2015a). A challenge facing nonexperimental research on the EITC is that major household surveys do not collect information on credit eligibility, receipt, or amount. Instead, EITC eligibility and amounts are simulated based on survey reports of age, family structure, earnings, income, and limited other information salient to tax liability, under the maintained assumption of 100 percent take-up conditional on eligibility. However, estimates using administrative Internal Revenue Service (IRS) data place actual take-up rates closer to 80 percent, and it is procyclical (Scholz 1994; Plueger 2009; Jones 2014). This suggests that estimates of antipoverty impacts from simulations assuming complete take-up may be overstated, especially in periods of economic decline. However, because the IRS pays some portion of EITC claims that are falsely made (Marcuss et al. 2014), simulated benefits received accounted for only two-thirds of aggregate EITC benefits paid (Meyer 2010), and thus survey simulations may understate the actual impact of the credit because they miss false credit claims. Furthermore, incorrect assignment of tax filing status and composition may lead to over- or under-estimates depending on whether families are systematically allocated close to or far away from the poverty line.

[1] See Hotz and Scholz (2003) and Nichols and Rothstein (2016) for comprehensive surveys of research on the EITC.

In this paper, we use a unique, internal Census dataset linking survey information from the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS) to administrative IRS tax data from Form 1040, W-2 wage statements, and the EITC return file to estimate how the distribution of model-based simulated EITC benefits compares to actual administrative EITC payments for the same individuals spanning tax years 2005-2014, and the subsequent effects on after-credit poverty rates. We construct six models of the EITC from the linked CPS ASEC-IRS data based on payment source and eligibility. The model using the IRS EITC recipient file, which records the actual EITC dollars paid to taxpayers, serves as our benchmark. This file provides estimates of the total dollar amounts of EITC credits circulating in the economy, regardless of whether those payments are correct. Restricting the latter model to those eligible provides an estimate of incorrect payments, and the ratio of those receiving IRS payment, conditional on eligibility, to those eligible who may or may not receive payment provides official estimates of EITC take-up (Pluger 2009; Jones 2014). This implies that a comparison of EITC payments from the two IRS models provides the effect of the missing counterfactual of those who are eligible but do not file for the credit.

Against the benchmark of actual EITC receipt, we focus on three EITC simulators: one produced by the Census Bureau as part of its annual release of the CPS ASEC (the CPS model), another produced by the National Bureau of Economic Research's TAXSIM model, and the third using Jon Bakija's tax model. Assessing the results of these tax simulators using survey data versus administrative data on a leading antipoverty program is important as the CPS ASEC is the source of official income and poverty statistics in the United States (Proctor et al. 2016), as well as the Census Bureau's research supplemental poverty measure (SPM) (Renwick and Fox 2016). Meanwhile, for the past 40 years TAXSIM has been the leading tax simulator used in the wider research community (Feenberg and Coutts 1993). Recently, researchers have also turned to Jon Bakija's tax model (Bakija 2009), which provides estimates that are closely correlated with TAXSIM's results when the inputs provided are sufficiently similar (Wheaton and Stevens 2016). We use the Bakija model on both survey-derived inputs and inputs derived from administrative data.[2]

With each tax model, we first compare weighted recipiency and benefit payment rates to the full internal IRS EITC file both overall and by tax filing status and number of qualifying children. Here we find that the CPS model deteriorates slightly over time (between our first and last year of data) in terms of matching the administrative totals of recipients and benefits paid, capturing only 61 percent of recipients and 57 percent of benefits paid in 2014, compared to 67 and 60 percent, respectively, in tax year 2006. The TAXSIM and Bakija models, however, are fairly stable over time, capturing around 75 percent of recipients and 70 percent of benefits paid in 2014 compared to around 79 and 68 percent in 2006. There is substantial heterogeneity in how well the inputs into the simulation models stack up to administrative totals across tax filing status and number of qualifying children, matching well to administrative totals for joint filers, but substantially less well for childless tax units.

We next examine the full distribution of EITC payments across both the individual tax unit and family income distributions. The former is informative as the EITC is targeted to tax filing units, and not families per se, while the latter is informative as it relates to how EITC

[2] Wheaton and Stevens (2016) compared how these tax models, as well as Urban Institute's TRIM model and the Bakija (2014) model, affect estimates of the SPM in tax year 2012, but they did not have access to administrative tax records in their analysis. In general, TAXSIM relies on a file transfer protocol to feed in the input file from a remote server and send back the output file in a return exchange. Clearly, this exchange would not be suitable for use with confidential data. In such cases, TAXSIM may be deployed using a local executable; however, while we had access to this executable when using the internal version of the ASEC, we do not have permission to use it with IRS-provided variables.

dollars flow to the families defined by the Census for official poverty statistics. Here we find that significantly more actual EITC payments flow to childless tax units than predicted by the tax simulators, and to those whose family income places them well above official poverty thresholds. However, those “incorrect” payments still flow to low-income tax units and thus appear target efficient at the taxpayer level, and some of the discrepancy at the family level is accounted for by multi-tax-unit families.

We then examine how many persons each EITC payment model lifts out of poverty. All three survey tax simulators provide comparable estimates of the antipoverty effect of the EITC in the full CPS ASEC sample; however, compared to the benchmark sample of IRS EITC payments circulating in the economy, each model overstates the number lifted out of poverty by as much as 1.8 million persons in recent years. Even though aggregate EITC payments in the CPS, TAXSIM, and Bakija models fall short of administrative totals, they tend to allocate those payments to families closer to the poverty line. The net effect is that the estimates derived from the survey-based tax simulators predict a much greater antipoverty bite for the EITC than found in administrative tax records, and only some of the overstatement is due to the use of survey income as inputs. However, when restricting the sample to a harmonized group of potential filers to make an “apples-to-apples” comparison, we find that the TAXSIM and Bakija models align closely with antipoverty estimates using actual EITC payments, but the CPS model continues to substantially overestimate the antipoverty effects of the EITC by over 700,000 persons. The assumption of 100 percent take up in the TAXSIM and Bakija models seems to balance out with the seemingly incorrect payments in the actual EITC recipient file, yielding comparable antipoverty effects. This does not translate in the CPS model, suggesting that the designation of tax filer, which differs between the full and harmonized samples, may not be as precise as in the *TAXSIM* or *Bakija* models.

II. Background on the EITC

The EITC was established in 1975 to incentivize work over welfare (“workfare”) by providing a refundable tax credit to families with qualifying children and low earnings, thereby creating a subsidy to market wages. The credit has three ranges—the “phase-in” or subsidy range where the credit amount increases at a fixed rate as earnings increase; the “plateau” range where the maximum credit is attained and held fixed; and the “phase-out” range where the credit is tapered away as earnings increase. The EITC was initially modest in size, largely offsetting payroll tax liability; however, it was expanded in both generosity and reach with the Tax Reform Act of 1986 and subsequently with the Omnibus Budget Reconciliation Acts (OBRA) of 1990 and 1993. OBRA90 differentiated tax units into those with one qualifying child versus two or more, and provided a more generous credit to those with two or more children. OBRA93 further expanded access to single individuals with no dependents, and also substantially increased the generosity of the credit with qualifying children such that by 1996 the EITC nearly reached the fourth decile of the married-couple income distribution and 150 percent of the median for female-headed families (Ventry 2000). A temporary, higher subsidy tier was added for families with three or more qualifying children with the American Recovery and Reinvestment Act (ARRA) of 2009, which was then made permanent in 2015. Table 1 summarizes the key parameters of the EITC in tax years 2005 and 2014, coinciding with the start and end of our sample. Across tax years the maximum credit is adjusted upward with inflation, though 2014 also contains the new category with a maximum subsidy rate of 45 percent and credit of $6,143.

From inception in 1975 through 2014 the number of EITC recipients grew over four-fold to 28.5 million families, with the average inflation-adjusted credit growing ten-fold to $2,395.[3]

[Table 1 here]

The growth in the credit has generated a voluminous research literature, including estimates of the antipoverty impact. While food stamps are estimated to be effective at reaching those in deep poverty (Tiehen, Jolliffe, and Smeeding 2015), the EITC is widely touted as the most successful program for low-income working families. As Hoynes and Patel (2017) note, the EITC can affect poverty mechanically via the credit amount on after-tax income, as well as behaviorally by affecting both the extensive and intensive margins of labor supply (Eissa and Leibman 1996; Meyer and Rosenbaum 2001; Neumark and Wascher 2001; Eissa and Hoynes 2004). Ziliak (2015a), using public release versions of the CPS ASEC and the CPS tax model, estimated that the EITC lifted about 4 million persons out of poverty (based on the official poverty line) per year in the decade prior to the Great Recession in 2008, and over 5 million at the peak of the recession. Hoynes and Patel, using the public CPS ASEC along with TAXSIM, estimated that the EITC lifted 3.4 million children in single-mother families out of poverty (the authors used the official poverty line but defined a broader resource definition that included some in-kind transfers and tax payments and credits). The Center on Budget and Policy Priorities (2016) estimated that 6.5 million people were lifted out of poverty in 2015 using the Census SPM and public CPS ASEC with the CPS tax model.[4]

A challenge facing research on the EITC is that none of the major household surveys in the U.S. collect information on credit receipt or amounts; more generally, they do not collect

[3] Brookings/Urban Tax Policy Center http://www.taxpolicycenter.org/statistics/eitc-recipients

[4] The larger estimate in the Center on Budget and Policy Priorities report results from the fact that the SPM poverty threshold is higher up the income distribution than the official line, and thus can capture a larger share of the EITC recipient population.

information on tax unit formation or tax deductions, and thus researchers either construct their own estimates of the EITC using program parameters or rely on publicly provided tax simulators such as the CPS tax model or TAXSIM (Ziliak 2015b). For the EITC the key inputs are earnings, interest income, age of filer, age and number of qualifying children, and filing status. The Census Bureau since the 1980s has used available survey information in the CPS ASEC to construct tax units and simulate federal, state, and payroll tax liability, including the EITC, and makes the output available in public release versions of the CPS ASEC. Each user of TAXSIM and the Bakija model, however, has to independently compile this information from the CPS ASEC or other survey prior to submitting the input to either simulator. While most surveys have enough information to create pointers to family and household relationships, these may not translate directly into tax filing units; that is, a family or household may contain multiple filing units. In addition, determining who does and who does not meet the "qualifying child" test is difficult. Children under age 19 must live with the filer at least six months and a day during the tax year, but the surveys do not typically record length of time in the household. This residency issue is further complicated by children of divorced parents who may spend equal time in each household. Children older than 18 years and younger than 24 may also be claimed as dependents if their primary activity is a full-time student, which is collected in some, but not all, household surveys. The net result is that the researcher must make assumptions on family relationships that may have direct impacts on the quality of tax estimates. The aim of our project is use a direct match of the CPS ASEC to administrative tax records to assess how well these models perform relative to actual payment in the antipoverty effect of the EITC.

III. Data

The data we use in this project derive from a joint statistical project agreement between the Census Bureau and the IRS. The Census Bureau receives tax records from the IRS to calculate and report on the take-up rate of the EITC, with the calculation of the denominator dependent upon survey data that is representative of the U.S. population. The survey data allow us to determine the members of the population who appear to be eligible, regardless of whether they file a Form 1040. The process of eligibility modeling and take-up calculation is reported in detail for tax year 2005 in Plueger (2009). The process, briefly described below, has changed somewhat in subsequent years, mainly in the refinement of income measurement.

The survey data used are yearly internal CPS ASEC files from 2006 to 2015, matched at the individual level for the corresponding tax year with the IRS data (that is, 2005 to 2014).[5] The tax data included in the project are Form 1040 individual income tax records, the EITC recipient file, the CP09/27 file (a record of taxpayers sent a notice from the IRS about their potential EITC eligibility), and Form W-2 wage and tax statements.

The records are made linkable using a process whereby individuals in each data set were given a unique, protected identification key, or PIK. When a Social Security Number (SSN) is available in a data set (such as all of the IRS records used in this project), the PIK is assigned based on SSN. Identifier placement is close to 100 percent in the case of administrative tax records with an SSN. The CPS ASEC stopped collecting SSNs as of the 2006 survey year, and thus personally identifiable information such as name, address, and date of birth is used in probabilistic matching against a reference file to assign PIKs (Wagner and Layne 2014). Personal information is then removed from each data set before it may be used for research

[5] The internal ASEC differs from the public-use ASEC primarily in terms of top-codes of earnings and income. For example, the internal top code of earnings is $1.09 million, while it is $250,000 in public versions.

purposes. For the IRS EITC estimation project, we also remove persons whose income and wage values were imputed in the CPS ASEC, as initial EITC eligibility determination, which uses only the survey data, is dependent on these values. We then reweight the data based on the probability that an observation received a PIK and did not have imputed or edited information.[6] We assess the impact of these sample choices throughout the paper.

A. Tax Simulator Models

The goal of this paper is to compare six methods of EITC eligibility and payment estimation—three based on survey data and tax simulators and three based on linked survey and direct tax data. The first survey-based approach is the method long used within Census to impute EITC receipt from the survey values alone (*CPS*).[7] This measure is part of a larger package of simulated tax variables that Census has provided to users of the internal- and public-use CPS ASEC, first developed in the 1980s and revised in survey year 2004 (U.S. Bureau of the Census 1993; O'Hara 2004). The CPS model first computes payroll tax liability for each person with earned income. It then constructs potential filing units (single, joint, head of household, dependent filer) based on marital status and household relationships (i.e., a household may have multiple tax filing units from related or unrelated subfamilies/individuals). These units are then statistically matched to the Statistics of Income (SOI) Public Use File to impute capital gains and itemized deductions, which are not collected in the CPS ASEC. This provides input to compute

[6] Bollinger et al. (Forthcoming) show that earnings nonresponse in the ASEC is over 30 percent on average and U-shaped across the (linked administrative) earnings distribution—highest in left and right tails—and not missing at random, meaning that reweighting the data after dropping imputed values helps, but will not completely remove bias from imputation. Hokayem et al. (2015) show that imputation in the ASEC leads to an official poverty rate that is too low by about 1 percentage point, and reweighting using an inverse probability weight fills about 70 percent of that gap.

[7] For simplicity, we refer to "CPS," "TAXSIM," "Bakija," and "Bakija admin," italicized, to refer to the estimated EITC payment derived from each simulator (with estimates greater than zero identifying someone as both eligible and paid, according to the model in question). The use of italics distinguishes the estimates from the simulators themselves. Any critique of these estimates should not be taken as critiques of the simulators, as the estimates are dependent on the quality of the input file.

initial federal adjusted gross income (AGI) and tax liability and credits, which in turn are used to construct state tax payments and credits, and then final federal taxes are computed using the estimated state tax payments as a deduction (for itemizers). Census appends to the CPS ASEC its estimates of federal, state, and payroll tax payments, the EITC credit amount, and estimated filing status and pointers to dependents. For our purposes, we focus on the EITC variable (eit_cred).

The second modeling strategy, widely used by many researchers, employs the NBER's TAXSIM model (*TAXSIM*).[8] This tax simulator uses up to 21 input variables derived from source data that reflect tax unit characteristics, including marital status, age of the primary taxpayer (and secondary if present), along with their wage and salary income, state of residence, number and ages of dependents (for calculation of EITC, child tax credit, etc.), and other taxable and nontaxable income, and potentially deductible expenses (home mortgage interest, property tax).[9] Similar to the CPS model, TAXSIM uses the SOI Public Use File to impute itemized deductions for each filing unit and compares these to the standard deduction for each filing unit to determine whether the taxpayer is assumed to itemize, along with an iterative procedure between federal and state tax payments. There are a possible 38 outputs provided by TAXSIM, with the federal EITC (v25) the output of focal interest to our project.

The third simulator we examine is the tax model developed by Jon Bakija (*Bakija*). His simulator differs from TAXSIM in its programming language (SAS versus FORTRAN) and in the larger number of separate variables that must be fed into the process as part of the input file,

---

[8] We use version 9 of TAXSIM that has been installed for research purposes with the internal ASEC at the Census Bureau. It is the same as Internet TAXSIM (v9) available at http://users.nber.org/~taxsim/taxsim-calc9/. There is also available to public users a new version, TAXSIM 27, that inputs 27 variables instead of the 21 used in TAXSIM 9.

[9] TAXSIM is traditionally used with survey data such as the ASEC, ACS, SIPP, PSID, and others, but is also employed to flesh out information from administrative records when full tax information is missing.

but it otherwise is similar. Comparisons between TAXSIM and the Bakija models indicate that, conditional on receiving comparable input files, their output files are highly correlated (>0.9 for most output variables considered in our analysis). This external validity suggests that the two simulators work very well, conditional on receiving an accurate input file from the user.

For *TAXSIM*, *Bakija*, and IRS eligibility (described below), we use the CPS ASEC to create tax units and assign dependents to filers. Rather than use Census-constructed input variables such as filing status, we construct our own.[10] The first task is to assign the heads and spouses (if applicable) for each potential filing unit identified by a unique ID based on household sequence number, family sequence number, family position, and family type. The head can be of the primary family, of a related subfamily, of an unrelated subfamily, or a primary individual. We also allow for dependent filers. We then construct a variable for the number of dependents based on age of the child and relationship to head, including those between ages 18 and 24 who are full-time students (and thus can be claimed as dependents for the EITC) as well as foster children. CPS ASEC observations are assigned as nonfilers if they are a dependent child, as single if they are unmarried and have no dependents, as head of household if they are unmarried and with dependents, and as joint filers if they are married with or without children. Wage and salary income is constructed from CPS ASEC variables, including farm and self-employment earnings, while the other taxable and nontaxable income sources are assigned to the taxpayer according to input requirements of TAXSIM or Bakija. Each primary and secondary taxpayer is run through the simulator, but the tax values of the primary taxpayer are the only ones retained to avoid double counting.

[10] A Stata DO file with the code that prepares the ASEC data for input into TAXSIM is available from the authors at https://sites.google.com/site/jamesziliak/Home/Research/. Our approach is an update of that made available by Judith Scott-Clayton to researchers using the ASEC as inputs to TAXSIM via the Stata interface. See http://users.nber.org/~taxsim/to-taxsim/cps/cps-clayton/.

B. Administrative Tax Models

We compare these simulators to three others that use survey data on the filing unit in combination with actual IRS tax records. The first, *IRS Eligible*, consists of tax units who are estimated to be IRS eligible based on information in the CPS ASEC supplemented by administrative records, regardless of whether they file a return or not. That is, we assume 100 percent take-up, using administrative IRS income amounts from the Form 1040 and W-2 for those who file and get paid, and a predicted amount based on program parameters, survey information, and W-2 information for those who do not file. We follow the method outlined in Plueger (2009), where the assignment of persons to tax units and the identification of filers and qualifying children is essentially the same as the method used for *TAXSIM* and *Bakija*. However, throughout the modeling process, eligibility is refined by substituting values from the matched 1040 data whenever available. We presume the reference person in each family to be the primary tax filer for the identified tax unit. In a later adjustment, the tax information on eligibility is transferred from the householder to the spouse if it was the spouse who filed. Variables on tax unit earnings, income, and dependent support requirements (which, in the case of the EITC, is based on where a child lived for the tax year) are first taken from the survey data. Then, values from the tax data, when available, are swapped in for the survey values, and eligibility is refined based on these new values.[11]

The second CPS ASEC-IRS model consists of those tax units in actual receipt of administrative EITC payment by the IRS regardless of eligibility, which we denote as (*IRS*

[11] Once weighted, filers identified in the CPS ASEC represent 94 percent of 1040 filers and 93 percent of W-2 observations. It should be noted that determinations of eligibility are based on assessments of the household roster as reported in the CPS ASEC, and not on audits. When we say tax units are paid regardless of eligibility, we do not imply that these tax units would be found noncompliant or fraudulent if subjected to an audit. The Appendix to this paper enters into a more in-depth analysis of household structure and incorrect payment by exploiting the panel aspects of the CPS ASEC.

Paid). The *IRS Paid* model is informative as it provides estimates of the total dollar amounts of EITC credits circulating in the economy, regardless of whether those payments are correct. Finally, we are interested in true EITC receipt for CPS ASEC respondents who are jointly IRS eligible and paid the EITC (*IRS Paid & Eligible*). A comparison of the *IRS Paid* and *IRS Paid & Eligible* sample provides an estimate of the amount of incorrect EITC payments made in a given year, while the ratio of the latter to the former provides an estimate of EITC take-up.

C. Sample Summary Statistics

Our six modeling strategies lead to the selection of two samples, each of which plays a part in the analysis. The first, "off the shelf" sample is the full sample of the CPS ASEC, which any researcher is able to use for his or her own EITC or poverty modeling—i.e., with this sample one is able to replicate official Census poverty rates. To be consistent with the official measure of poverty—which is family-based and aggregates all primary and related subfamily members—in this sample we consider all heads of families to be potential filers, whether some of the income inputs are imputed or not. As such, we use the Census-provided CPS ASEC supplement weight. The four tax estimates that can be examined within this sample are *CPS*, *TAXSIM*, and *Bakija*, along with *IRS Paid*. This sample represents a population of 1.3 billion tax filers over the 10-year period. The second sample is a harmonized sample, where we select those heads of tax units who are identified as potential filers in the Census-IRS modeling strategy. This approach removes a potential filer from the full sample when income or earnings reports are imputed or if the potential filer did not receive a PIK. It also categorizes, when the data are reweighted, more persons as filers who were not originally identified in the full poverty sample—i.e., the harmonized sample adds in tax units whether they are the primary family head or head of a tax-filing unit within the larger family. We adjusted the CPS ASEC person weights to account for sample selection into the harmonized sample, leading to a population of 1.5 billion filer-years. The harmonized sample allows us to also examine *IRS Eligible* and *IRS Paid & Eligible.*

[Table 2 and 3]

Table 2 contains selected weighted summary statistics from the full linked CPS ASEC-IRS sample, followed by the harmonized sample in Table 3. In each table, the upper panel presents statistics at the individual tax filer level pooled across the 2005–2014 tax years, while the bottom panel is aggregated to the family level. All income amounts are in real 2009 dollars using the Personal Consumption Expenditure Deflator.[12] Table 2 shows that across all sample years, 13 percent of persons are estimated to be EITC eligible from the *CPS* tax model, and 14 percent from the *TAXSIM* or *Bakija* models. The average credit amount was $267, $271, and $277, respectively. In contrast, only 11 percent of the potential filers identified in the full sample actually receive EITC (*IRS Paid)*, and on average they receive about $253. Comparing these averages to Table 3, the situation flips, with the probability of eligibility using *CPS*, *TAXSIM*, and *Bakija* is around 12 to 13 percent, and 15 percent of the sample is actually paid EITC. Credit amounts average $244, $245, and $248, respectively, for *CPS, TAXSIM,* and *Bakija,* while *IRS Paid* averages $333. Taken together, the tables point to a difference in how filers are identified between the full and harmonized sample. The larger value of *IRS Paid* in Table 3 also suggests that more persons receive the credit than are eligible (14 percent), and the size of the credit received ($333) is larger than would be expected based on IRS eligibility criteria ($250).

[Table 4 here]

[12] Appendix Table B-3 in the 2017 Economic Report of the President, https://obamawhitehouse.archives.gov/sites/default/files/docs/appendix_b-statistical_tables_relating_to_income_employment_and_production_2017.pdf

Table 4 demonstrates how values of EITC differ dependent on the sample chosen and the source of the EITC estimate. A CPS ASEC observation may be a respondent or a non-respondent, and may or may not receive an identifier. Table 4 reports the *Bakija* estimate for each of the mutually exclusive groups defined by their respondent-plus-PIK category and by quintiles[13] of CPS ASEC adjusted gross income (truncated to less than $60,000 in 2009 dollars). For the categories with an identifier, we also provide the value of *IRS Eligible.* All of the income and earnings concepts that feed into the *Bakija* estimate for the non-respondents derive from a hot-deck procedure that swaps these values in from respondents who look similar on a variety of characteristics. The *Bakija* estimates differ substantially over the categories of respondent. The "non-respondent/PIK" group is assigned *Bakija* estimates that are roughly $100-$170 more than the "respondent/PIK" group for the lower two income bins, and values about $20-$30 less for the top two bins. The *IRS Eligible* values show a wider difference between these two groups at each point in the distribution—this is especially apparent in the third quintile, where the difference in *IRS Eligible* is about $230. Such differences highlight the impact that imputation has on the different estimates and samples we assess throughout the paper.

IV. Comparing Alternative EITC Payment Methods

We begin by establishing that the EITC recipient file submitted to Census aligns closely to published aggregates from the IRS's SOI, and thus can serve as the benchmark standard to measure the alternative modeling strategies. Table 5 contains both the numbers of recipients (in millions) and the benefits paid (in millions of nominal dollars) for tax year 2006 as reported in SOI public reports, as well as comparable numbers from the internal EITC recipient file submitted by IRS to Census annually for estimation of EITC eligibility and take-up. As in Meyer

[13] The first quintile reported a zero credit in all cases.

(2010) we present the statistics broken down by number of qualifying children, and also use tax year 2006 as in his study. The internal file covers 97 percent of EITC dollars and 100 percent of recipients reported publicly. There are some differences by number of children, with the internal file finding a slightly higher proportion of EITC recipients with zero children. This slight discrepancy may be due to amended filings, which could differ in the two sources depending on the vintage of the source.

[Table 5 here]

A. Mean Differences

Having established that the internal EITC recipient file compares very favorably to published aggregate totals, we next use that internal file as the benchmark to compare against the six EITC modeling approaches. Tables 6 and 7 compare each EITC estimate, using the harmonized sample reweighted to reflect the population, to the numbers in the internal EITC recipient file for tax years 2006 and 2014, respectively. Characteristics in the recipient file (filing status and number of children claimed for EITC) are compared with the corresponding characteristic calculated using the IRS Eligibility estimation strategy. In looking at the overall ratio between the internal recipient file and the weighted estimates in the bottom two rows of Tables 5 and 6, *CPS* covers about 60 percent of EITC dollars and 67 percent of EITC recipients in 2006, but only 57 percent and 61 percent, respectively, in 2014 (Table 7). *TAXSIM* captures 68 percent of dollars and 79 percent of recipients in 2006, and 70 percent and 75 percent, respectively, in 2014. *Bakija* covers about 69 percent of benefits and 74 percent of dollars in 2006, and 71 percent and 72 percent, respectively, in 2014.

[Tables 6 and 7 here]

At first glance, it appears that these coverage estimates from the tax simulators fall short relative to those obtained among who actually receive the EITC: in the *IRS Paid* model where we compare actual receipt against the internal recipient file we cover 97 percent of 2006 dollars and 92 percent of recipients (these are 94 percent and 87 percent in 2014). This difference is driven by the fact that many EITC recipients do not in fact appear to qualify. Indeed, when we restrict the *IRS Paid* model to those who are eligible (the last column in Tables 6 and 7), we find that the estimates from the survey tax simulators are much closer in terms of overall coverage (although the potential exists that the persons covered are not the same).

While these results suggest that, overall, we capture a substantial portion of the paid population using the linked the data, Tables 6 and 7 show that the quality of the estimates vary by filer characteristic. For example, in the *IRS Paid* column we calculate between 14 (Table 7) and 32 (Table 6) percent more recipients with zero qualifying children than the internal file reports. This is due to persons who claim one or more children on their 1040 but do not appear to have a qualifying child according to the survey data. However, much of the heterogeneity arises from incorrect payments. Every estimate over-reports the dollars going to filers with zero children. While this category represents only three percent of total EITC expenditures from the recipient file, each estimate reports between 1.5 and seven times more dollars accruing to this category.

A look at the *IRS Paid & Eligible* ratio makes clear that many of those who are incorrectly paid have inappropriately claimed one or more children, which accounts for the discrepancy between the recipient file and the estimates. The discrepancy is not entirely explained, however: we still find 1.5 times the dollar values going to recipients with zero children even when limiting the comparison to those deemed eligible.

[Figures 1 and 2 here]

In Figures 1 and 2 we examine how well the modeling strategies translate into average credit amounts in each year. Figure 1 shows the harmonized-sample inflation-adjusted mean values. In each year, the mean values from *CPS*, *TAXSIM*, *Bakija*, and *IRS Eligible* series are each quite close; however, they all fall short of the actual paid amount from the recipient file (*IRS Paid*) by at least $75-$100 depending on year. All models capture the increase in EITC generosity with the ARRA expansion in 2009, and track paid amounts. Figure 2 shows the same series, but restricted to those deemed by the Census-IRS EITC project to be eligible for the credit. In most years, *CPS*, *Bakija*, and *TAXSIM* are very similar. Notably, because we condition on eligibility, the gap of approximately $200 to $250 between *IRS Eligible* and *IRS Paid* now reflects the difference in average credit value induced by imperfect take-up—i.e., those eligible but failing to take up benefits would receive more on average than those that did.

B. Distributional Differences

We move beyond mean outcomes to examine the value of estimated EITC payments across the income distribution. For each estimate, we calculate the average value of the credit over bins of individual income, where the income measure is that used in estimating the credit value. In other words, *IRS Paid* and *IRS Eligible* are averaged over bins of IRS 1040 AGI, while *CPS*, *TAXSIM*, and *Bakija* are averaged over bins of CPS ASEC individual income. We use the harmonized sample to include the *IRS Paid* and *IRS Eligible* estimates.[14]

[Figure 3 here]

We graph the estimates using survey data alone in the top panel of Figure 3. All three simulator estimates are very close to one another and trace out the general shape of the EITC

[14] Recall that while earnings are necessary to qualify for the EITC, actual eligibility and credit amounts are determined by both earnings and AGI.

schedule. There is a drop to zero in credit receipt at the point in the income distribution where one would expect to see it (around $45,000). We next gray out these estimates, in the bottom panel, and trace out the IRS-derived estimates, including *IRS Paid*, *IRS Eligible,* and an estimate from the Bakija model that uses earnings and income information from the IRS, but does not otherwise adjust the input data (*Bakija admin)*. Compared with estimates derived from administrative records, the survey-derived estimates are right skewed, having more mass in the plateau region of the EITC. Meanwhile, while the *IRS Eligible* estimate drops to zero at the same point in the distribution as the survey-derived estimates, *IRS Paid* and the *Bakija admin* have considerable mass beyond that point. Moreover, *IRS Paid* lies above all other estimates over much of the income distribution. A clear finding of this exercise is that the difference between *IRS Paid* and *IRS Eligible* is widest at low levels of individual income, indicating that the payment of EITC is targeted to low-income people regardless of whether they are strictly eligible according to other parameters.[15]

[Figure 4 here]

Figure 4 breaks out the distributions from Figure 3 by number of children claimed and filing status. These variables combine information from both the survey and IRS data to generate the most likely claiming status for each tax unit. For ease of presentation, and because of the comparability of estimates, in this set of figures *CPS* stands in for all of the survey-based estimators. While the survey estimators come close to both *IRS Eligible* and *IRS Paid* for tax units with two or more children and for joint filers, the same cannot be said for units without children and with single filers. Our estimate using *Bakija admin* lies between the survey-based eligibility estimates and the IRS estimates, which provides some evidence that adjusting earnings

[15] This is not to deny that there are clearly tax filers beyond the $45,000 mark receiving credit dollars, apparently incorrectly.

and income using the administrative records only partially accounts for eligibility differences when comparing the tax and survey information.

[Figure 5 here]

Finally, it is important to assess how average credits are distributed over the variable that determines poverty status: family income, which is a survey-based measure collecting all sources of income for related individuals in a family, including subfamilies. In Figure 5, we graph out *CPS* (which again stands in for all of the survey-based estimates), *IRS Paid*, and *Bakija admin* in bins of family income from the CPS ASEC in the top panel and by income-to-needs in the bottom panel.[16] Here, we see that from the family-income perspective, the right skewness of the survey-based estimators is more pronounced than it is for individual income in Figure 3, highlighting how differences in income definitions lead to different conclusions about EITC targeting. Together, the evidence on average EITC vis à vis individual tax unit and family income indicates that true EITC Eligible dollars may disproportionately benefit *families* at higher income levels, while EITC dollars that look improper are still "targeted" primarily to individual *tax units* who are in or near poverty.

[Figure 6 here]

One possible reason for the apparent greater target efficiency at the tax unit income level versus the family income level is the possibility of multiple tax units within a given family. That is, a family may consist of more than one tax filing unit that is independently EITC eligible, but at the higher-level aggregation of the family would not be eligible if all those tax units filed a combined return. To examine this possibility, in Figure 6 we present average EITC amounts by bins of family income separately for single tax filing units and multiple tax filing units. It is clear

[16] Income-to-needs is computed by taking the ratio of family income to family-size specific poverty thresholds. Income-to-needs less than 1 indicates living in poverty.

from Figure 6 that the density in the multi-filer panel is thicker at higher income levels, and does not taper off as quickly after $50,000 in family income as in the single filer case.[17]

V. Anti-Poverty Effects of the EITC Comparing Alternative EITC Payment Methods

We next examine how the alternative modeling strategies align in terms of the number of persons lifted above the family-size specific official poverty threshold each tax year by the EITC. Here we conduct two exercises: first, we examine the "off the shelf" type of calculation by the typical user of the full-sample CPS ASEC and either the CPS, TAXSIM, or Bakija tax models. We then compare these estimates directly to our benchmark value, which is *IRS Paid*; second, we restrict attention to the harmonized sample used in the earlier analysis in order to conduct more of an "apples-to-apples" comparison of the six modeling strategies. It is only the latter exercise where we are able to report the *IRS Eligible and IRS Paid & Eligible* series. These two exercises are informative for different purposes: the first allows us to use the official definition of poverty calculated from the full CPS ASEC and examine the impact of true EITC receipt on that rate; the second allows us to compare estimation strategies over a specific set of potential filers.

[Table 8 here]

Table 8 contains the estimated poverty rates, where the top panel A is for the full CPS ASEC and the bottom sample B is for the reweighted harmonized sample. The column labeled ASEC in panel A replicates the official poverty rate as reported in the annual P60 report (Proctor et al. 2016), while the remaining fours columns contain the corresponding estimated poverty rate after adding in the EITC from *CPS, TAXSIM*, *Bakija*, and *IRS Paid*.[18] Panel A shows that the

[17] Jones and O'Hara (2017) show that multifiler households exhibit some strategic behavior in claiming EITC, using the movement of children across filers to increase EITC receipt.

[18] Historic time series of the poverty rate is available at https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-people.html

poverty rate is everywhere below the official rate once the credit is added in as a family resource, but the antipoverty effect is much larger for the survey tax simulators compared to *IRS Paid*. This is seen transparently in Figure 7, which depicts the total number of persons lifted out of poverty in panel A and the corresponding number of children in panel B. The three tax simulators predict that the EITC lifts 4.8 million persons out of poverty in an average year (about 2.75 million children), while *IRS Paid* lifts 3.2 million persons, or 33 percent fewer, and this gap was exacerbated over the sample period.

[Figures 7-8 here]

Figure 8 conducts the same exercise as Figure 7 but now we examine the total number of persons lifted above three alternative thresholds of 50% of the poverty line (a standard measure of “deep poverty”), 75% of the poverty line, and 125% of the poverty. There we see that the tax simulators lift comparable numbers out of deep poverty, but the anti-poverty effect of the EITC is more prominent the further up the income distribution for the survey-based tax simulators compared to actual administrative EITC payments. This suggests that the simulators assumption of 100% take-up is likely picking up more persons close to the poverty line.

[Figure 9 here]

To examine this further, in the bottom panel B of Table 7 and in Figure 9, we restrict the analysis to the common, harmonized sample without imputations, and reweight. We now find that *TAXSIM* and *Bakija* offer comparable estimates of poverty and the number lifted above the poverty line by the EITC as the *IRS Paid* and *IRS Eligible* samples. However, the *CPS* results in an estimate of the number lifted out of poverty about 700,000 persons higher compared to *TAXSIM*, *Bakija*, *IRS Paid*, and *IRS Eligible*. The latter, however, are themselves 600,000 - 700,000 higher than that predicted by the *IRS Paid & Eligible* sample, meaning that even though some of the EITC dollars in the *IRS Paid* model reflect incorrect payments, they are generally target efficient at lifting many out of poverty.

Comparing Figures 7 and 9 suggests that the antipoverty impacts of the EITC from the popular tax simulators are too high, and this is especially so for *CPS*, which in recent years predicts nearly one million more persons lifted out of poverty than *TAXSIM*. In our harmonized sample, there is an issue of designation of tax filer in the CPS model that does not affect the *TAXSIM* or *Bakija* models. The latter two models do well when comparing against the paid amount regardless of eligibility, since the 100 percent take-up assumption seems to balance out against incorrect payment in the *IRS Paid* model. The implication is that the *CPS* model is not identifying filers as effectively the other two simulators.

VI. Conclusion

The EITC is an effective antipoverty program for non-seniors in the United States. However, the program's antipoverty impact has, in the past, been calculated based on simulated EITC benefits using either the Census Bureau's tax module or from tax simulators such as the NBER TAXSIM or Bakija module. The simulation programs utilize researcher-provided model-based assumptions on who is and who is not eligible for the EITC based on survey values, and conditional on eligibility, assume that participation is 100 percent.

As shown in analyses of linked survey and administrative records, EITC take-up is considerably less than 100 percent. Claims regarding the impact of the program on measures of poverty may be overstated if the credit estimation relies solely on survey data and if take-up and incorrect payment are not accounted for. We find that, when using the full CPS ASEC, EITC benefits estimated from the CPS tax module and their impact on poverty are overestimated compared with actual benefit payments. While TAXSIM and Jon Bakija's model, when used with survey data alone, appear to perform better than the CPS tax module in the full CPS ASEC, a comparison between the TAXSIM and Bakija models with actual EITC payment among those eligible for a harmonized sample of filers indicates that the simulators perform well largely due to overestimation of take-up and underestimation of incorrect payment. Analysis of the average credit values over bins of survey and administrative income suggests that “true” EITC eligibility targets families at higher values of family income than do the survey-based estimates. Meanwhile the average value of EITC dollars actually paid is greater than the value of each eligibility estimate over the entire individual income distribution, suggesting that although some EITC is incorrectly paid to ineligibles, much of those dollars target individual tax units that are in poverty.

References

Bakija, Jon. 2014. “Documentation for a Comprehensive Historical US Federal and State Income Tax Calculator Program.” Williamstown, MA: Williams College.

Bollinger, Christopher R., Barry Hirsch, Charles Hokayem, and James P. Ziliak. Forthcoming. “Trouble in the Tails? What We Know About Earnings Nonresponse Thirty Years After Lillard, Smith, and Welch.” *Journal of Political Economy*.

Center on Budget and Policy Priorities. 2016. “Policy Basics: The Earned Income Tax Credit.”

Chetty, Raj, John Freidman, and Emmanuel Saez. 2013. “Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings,” *The American Economic Review* 103(7): 2683-2721.

Falk, Gene, and Margot Crandall-Hollick. 2016. “The Earned Income Tax Credit (EITC): An Overview,” Congressional Research Service Report 7-5700.

Eissa, Nada, and Jeffrey Leibman. 1996. “Labor Supply Response to the Earned Income Tax Credit.” *The Quarterly Journal of Economics* 111(2): 605-637.

Eissa, Nada, and Hilary Hoynes. 2004. “Taxes and the Labor Market Participation of Married Couples: The Earned Income Tax Credit.” *Journal of Public Economics* 88(9-10): 1931-1958.

Feenberg, Daniel, and Elisabeth Coutts. 1993. “An Introduction to the TAXSIM Model,” *Journal of Policy Analysis and Management*, 12(1): 189-194.

Hokayem, Charles, Christopher R. Bollinger, and James P. Ziliak. 2015. “The Role of CPS Nonresponse in the Measurement of Poverty.” *Journal of the American Statistical Association* 110(511): 935-945.

Hotz, V. Joseph, and John Karl Scholz. 2003. “The Earned Income Tax Credit,” In *Means-Tested Transfer Programs in the United States*, R.A. Moffitt (ed), Chicago: NBER and University of Chicago Press, 141-197.

Hoynes, Hilary, and Ankur Patel. 2017. “Effective Policy for Reducing Poverty and Inequality? The Earned Income Tax Credit and the Distribution of Income,” *Journal of Human Resources*, doi:10.3368/jhr.53.4.1115.7494R1.

Jones, Maggie R. 2014. “Changes in EITC Eligibility and Participation, 2005-2009,” U.S. Census Bureau, Center for Administrative Records Research and Applications, CARRA-WP-2014-04.

Jones, Maggie R., and Amy B. O’Hara. 2016. “Do doubled-up families minimize household-level tax burden?" *National Tax Journal* 69(3): 613-640.

Jones, Maggie R. 2018. Tax Preparers, Refund Anticipation Products, and EITC Noncompliance. Unpublished manuscript.

Marcuss, Rosemary, Alain Dubois, Janice Hedemann, Mary-Helen Risler, and Kara Leibel. 2014. *Compliance Estimates for the Earned Income Tax Credit Claimed on 2006-2008 Returns*. Internal Revenue Service, Publication 5162, Washington DC.

Meyer, Bruce D. 2010. "The Effects of the Earned Income Tax Credit and Recent Reforms," In *Tax Policy and the Economy*, 24(1): 153-180.

Meyer, Bruce and Dan Rosenbaum. 2001. "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *Quarterly Journal of Economics* 116(3): 1063-1114.

Neumark, David, and William Wascher. 2001. "Using the EITC to Help Poor Families: New Evidence and a Comparison with the Minimum Wage," *National Tax Journal* 54(2): 281-317.

Nichols, Austin, and Jesse Rothstein. 2016. "The Earned Income Tax Credit," In *Economics of Means-Tested Transfer Programs in the United States, Volume 1*, Robert A. Moffitt (ed.), Chicago, IL: The University of Chicago Press, 137-218.

O'Hara, Amy. 2004. "New Methods for Simulating CPS Taxes," U.S. Census Bureau, SEHSD Working Paper 2004-08.

Plueger, Dean. 2009. "Earned Income Tax Credit Participation Rate for Tax Year 2015," Internal Revenue Service Working Paper, https://www.irs.gov/pub/irs-soi/09resconeitcpart.pdf.

Proctor, Bernadette D., Jessica L. Semega, and Melissa A. Kollar. 2016. "Income and Poverty in the United States," U.S. Census Bureau, Current Population Reports, P60-256.

Renwick, Trudi, and Liana Fox. 2016. "The Supplemental Poverty Measure: 2015," U.S. Census Bureau, Current Population Reports, P60-258.

Scholz, John Karl. 1994. "The Earned Income Tax Credit: Participation, Compliance, and Antipoverty Effectiveness," *National Tax Journal* 47(1): 63-87.

Tiehen, Laura, Dean Jolliffe, and Timothy Smeeding. 2015. "The Effect of SNAP on Poverty," In *SNAP Matters: How Food Stamps Affect Health and Well-Being*, J. Bartfeld, C. Gundersen, T. Smeeding, and J. Ziliak (eds), Stanford: Stanford University Press, 49-73.

U.S. Bureau of the Census. 1993. *Measuring the Effect of Benefits and Taxes on Income and Poverty: 1992*, Current Population Reports, Series P60-186RD, U.S. Government Printing Office, Washington, DC.

Ventry, Dennis. 2000. "The Collision of Tax and Welfare Politics: The Political History of the Earned Income Tax Credit, 1969-99". *National Tax Journal* 53(4, Part 2): 983–1026.

Wagner, Deborah, and Mary Layne. 2014. “The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications’ (CARRA) Record Linkage Software,” CARRA Working Paper Series #2014-01, U.S. Census Bureau.

Wheaton, Laura, and Kathryn Stevens. 2016. “The Effect of Different Tax Calculators on the Supplemental Poverty Measure,” The Urban Institute.

Ziliak, James P. 2015a. “Recent Developments in Antipoverty Policies in the United States.” In *Social Policies in an Age of Austerity*, J.K. Scholz, H. Moon, and S. Lee (editors), Cheltenham, UK: Edward Elgar Publishing, 235-262.

Ziliak, James P. 2015b. “Income, Program Participation, Poverty, and Financial Vulnerability: Research and Data Needs,” *Journal of Economic and Social Measurement* 40(1-4): 27-68.

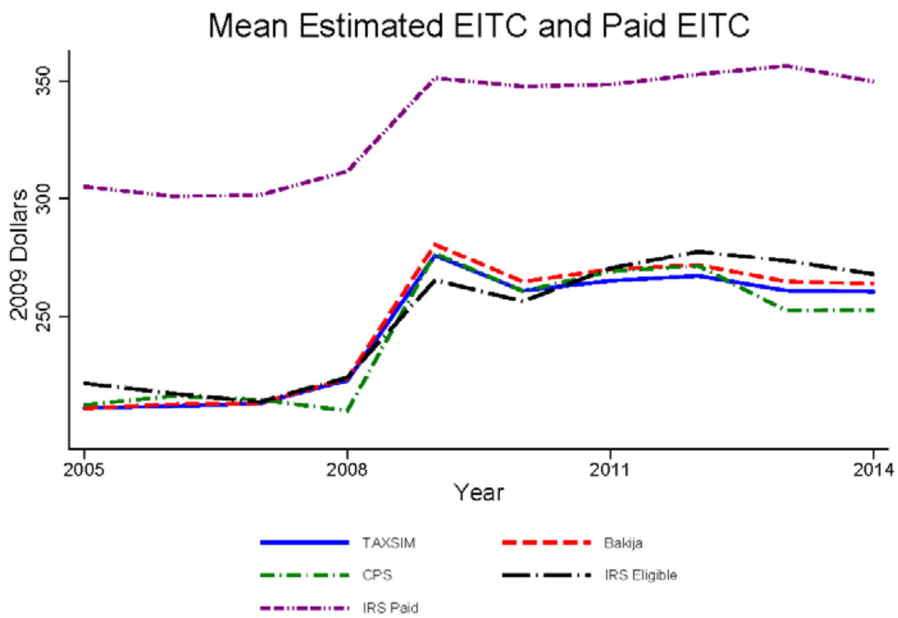


Figure 1. Mean estimated EITC from TAXSIM, Bakija, and Census CPS tax models, and estimated and paid EITC from combined CPS ASEC and IRS data. Full harmonized sample (see text for details). Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2014.

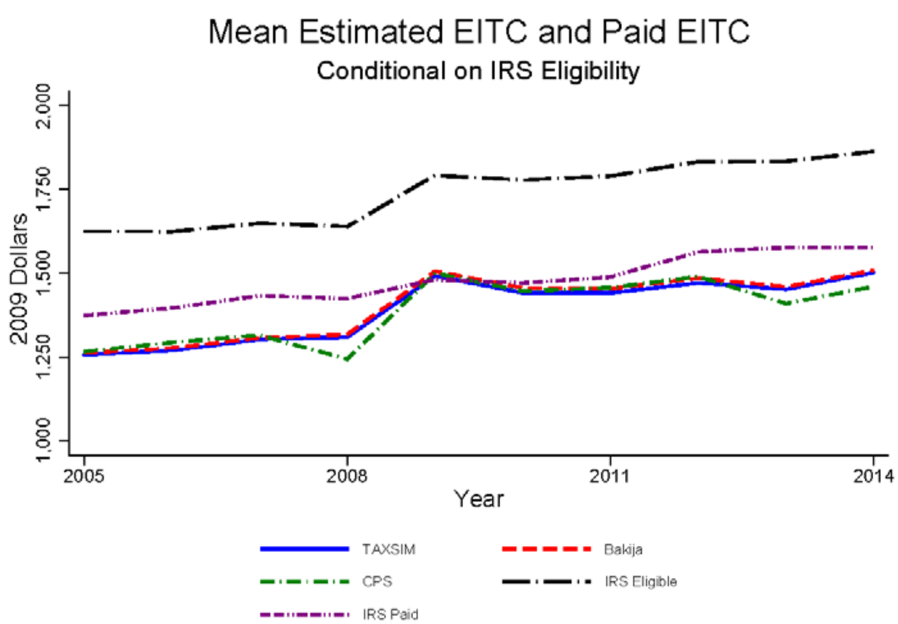


Figure 2. Mean estimated EITC from TAXSIM, Bakija, and Census CPS tax models, and estimated and paid EITC from combined CPS ASEC and IRS data. Full harmonized sample restricted to potential taxpayers who are EITC eligible according to official estimation (see text for details). Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015.

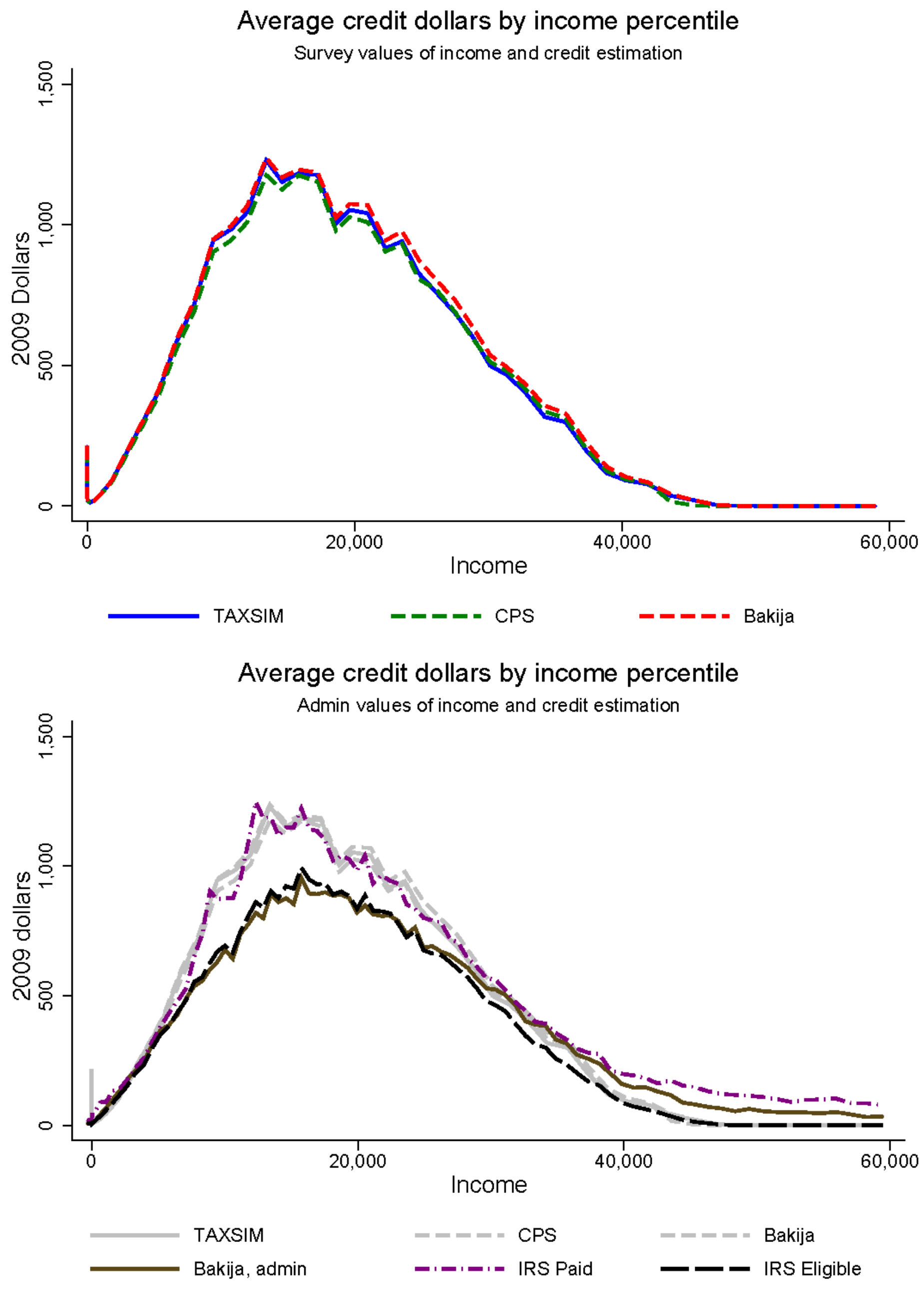


Figure 3. Top: Average credit dollars by income percentile for estimates using CPS-derived values of income. Bottom: Values of estimates derived from using administrative tax records for both income and EITC receipt. Full harmonized sample (see text for details). Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015.

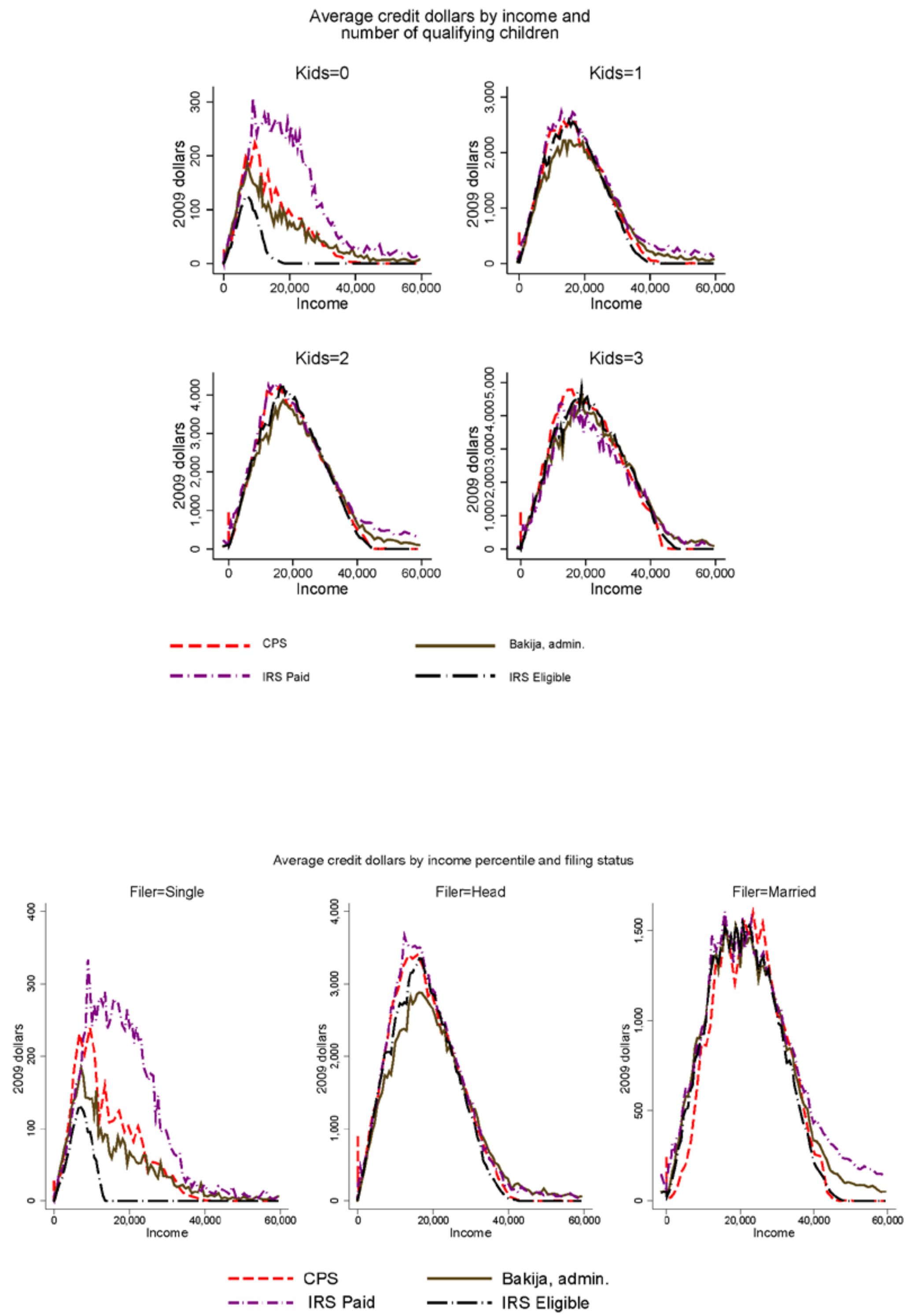


Figure 4. Top: Average credit dollars by bins of income, broken out by number of children in household. Bottom: The same, broken out by filing status. Full harmonized sample (see text for details). Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015.

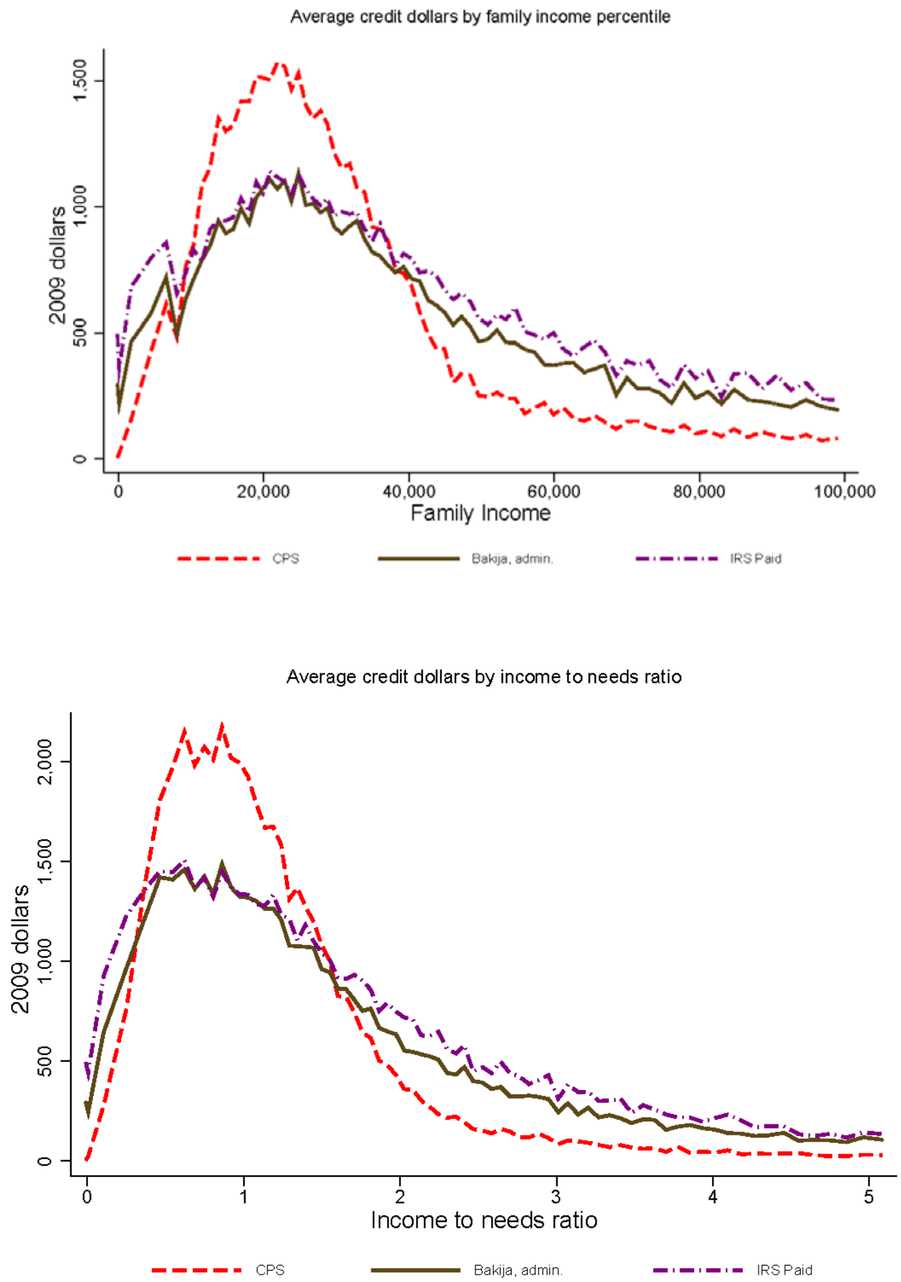


Figure 5. Average credit dollars by bins of family income. Here, the CPS Eligible estimate stands in for all survey-based estimates, as these were closely aligned. The line labeled “Bakija Eligible, admin.” gives the estimate using the same input file as the survey version of the estimate with only income and earnings replaced with tax-record variables. Full poverty sample. Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015.

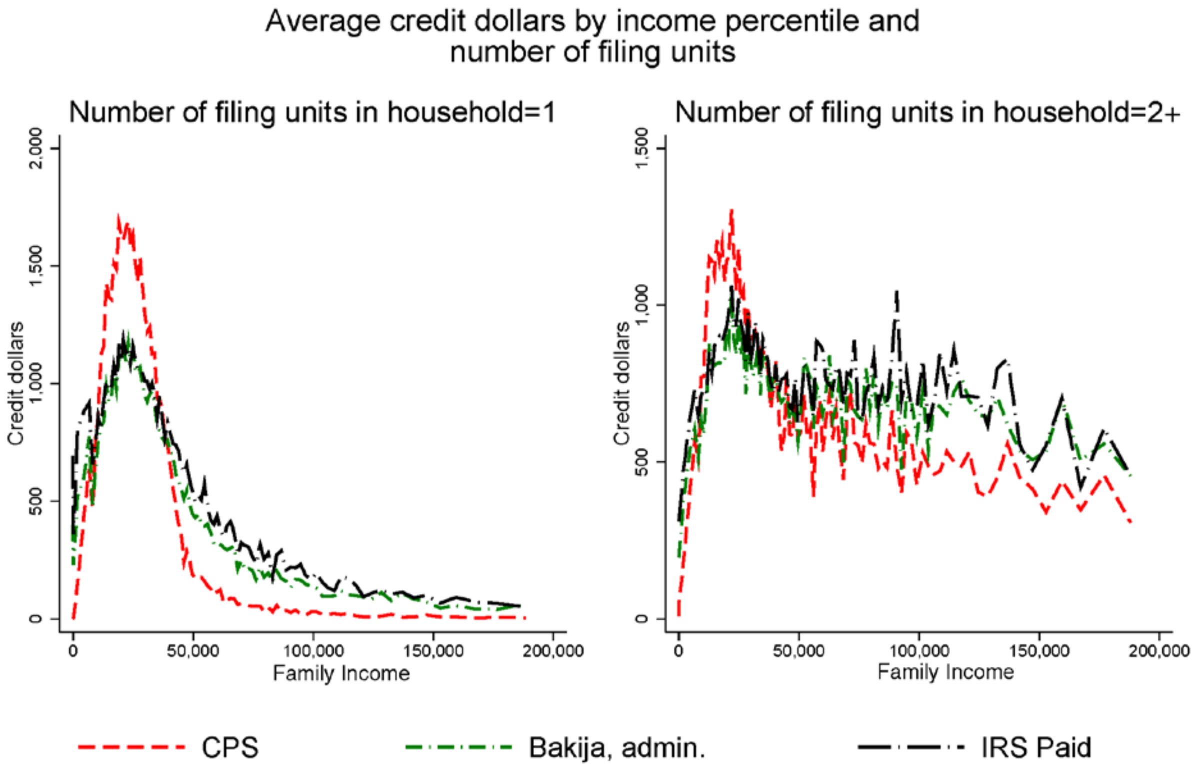
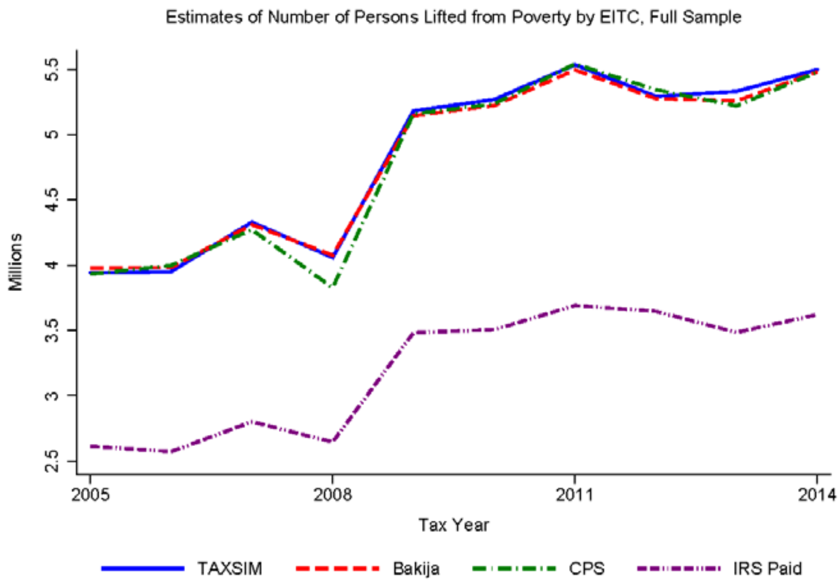


Figure 6. Average credit dollars by bins of family income, by number of tax filing units. Here, the CPS Eligible estimate stands in for all survey-based estimates, as these were closely aligned. The line labeled "Bakija Eligible, admin." gives the estimate using the same input file as the survey version of the estimate with only income and earnings replaced with tax-record variables. Full poverty sample. Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015.

A. All Persons



B. Children

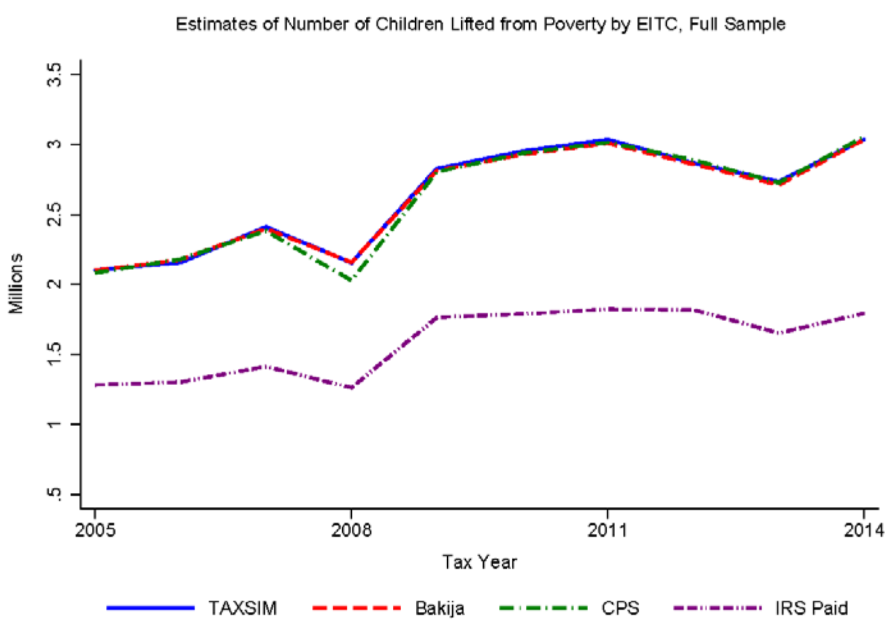


Figure 7. Estimates of the number of persons lifted from poverty by EITC using the full sample and original survey weights. Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015.

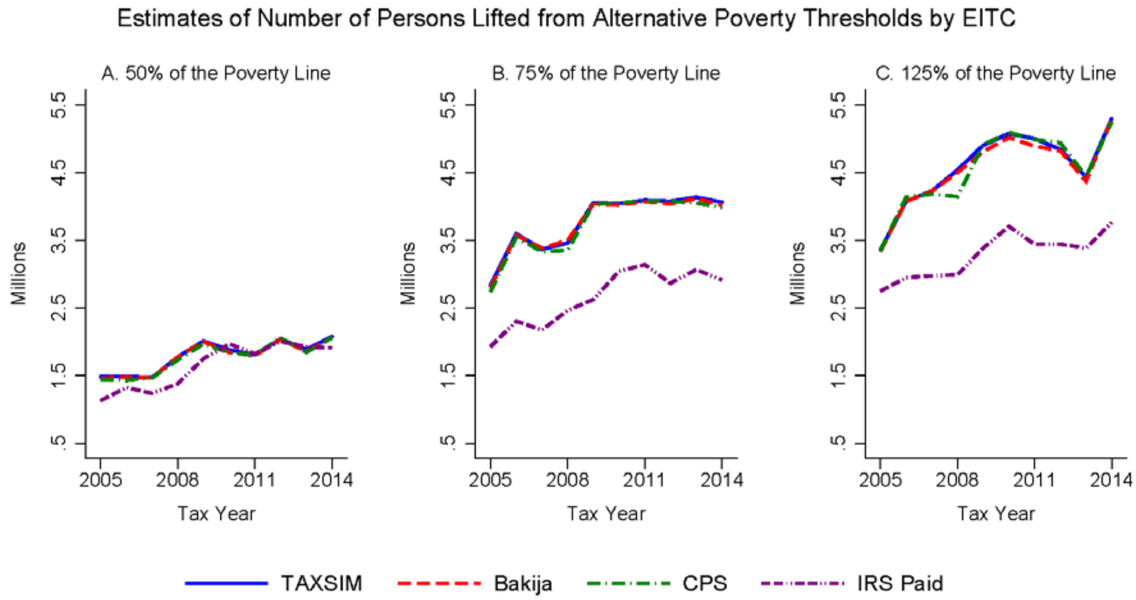


Figure 8. Estimates of the anti-poverty effect of EITC at different poverty thresholds. Full sample. Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015.

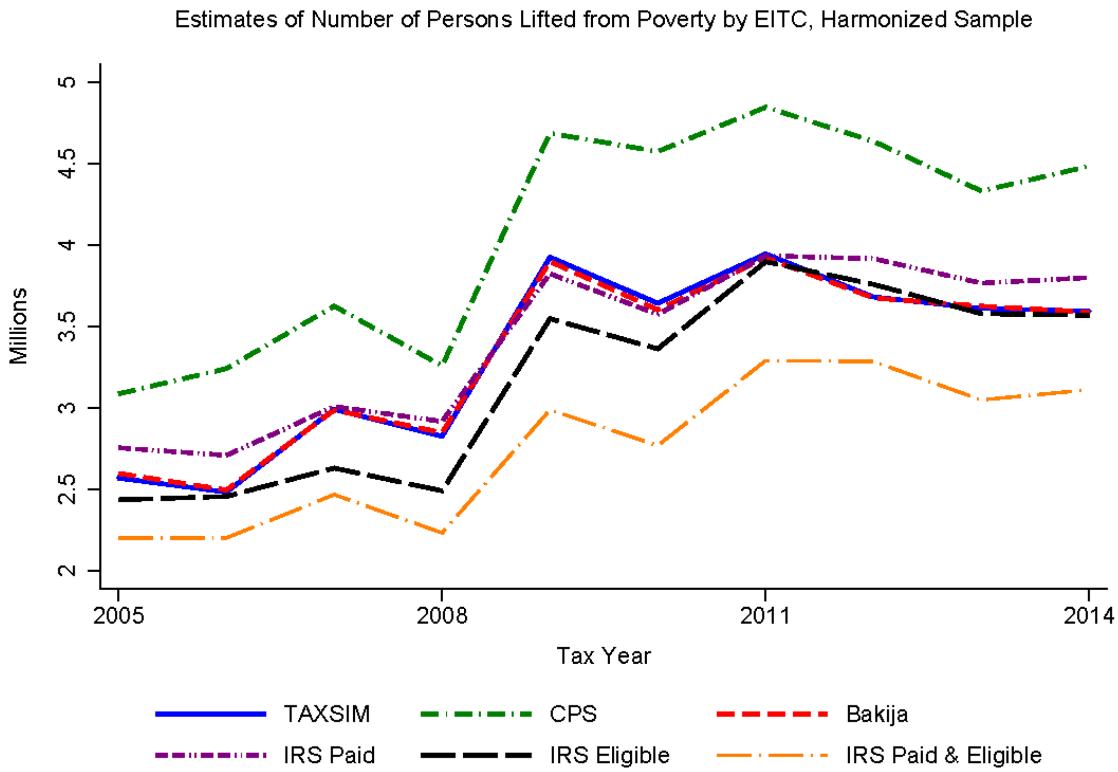


Figure 9. Estimates of the number of persons lifted from poverty by EITC using the harmonized sample and weights adjusted for sample selection. Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015.

Table 1. Earned Income Tax Credit Parameters, 2005 and 2014 Tax Years

Qualifying Children	Credit rate (percent)	Minimum income for maximum credit	Maximum credit	Phaseout rate (percent)	Phaseout range: Beginning income	Phaseout range: Ending income
2005						
No children	7.65	5,220	399	7.65	6,530	11,750
One child	34	7,830	2,662	15.98	14,370	31,030
Two children	40	11,000	4,400	21.06	14,370	35,263
2014						
No children	7.65	6,480	496	7.65	8,110	14,590
One child	34	9,720	3,305	15.98	17,830	38,511
Two children	40	13,650	5,460	21.06	17,830	43,756
Three children	45	13,650	6,143	21.06	17,830	46,997

Source: Brookings/Urban Tax Policy Center http://www.taxpolicycenter.org/statistics/eitc-parameters

Table 2. Summary Statistics for Linked CPS ASEC-IRS Data, Full Sample of Potential Tax Filers

Full sample	Mean	Standard Err.	95% Confidence Interval	
Reweighted (Obs.= 842 thou., N=1,378 mil.)				
EITC Receiver (*CPS*)	0.13	0.001	0.13	0.13
EITC Receiver (*TAXSIM*)	0.14	0.001	0.14	0.14
EITC Receiver (*Bakija*)	0.14	0.001	0.14	0.14
EITC Receiver (*IRS Paid*)	0.11	0.001	0.11	0.12
EITC Amount (*CPS*)	267.40	1.327	264.70	270.00
EITC Amount (*TAXSIM*)	271.30	1.323	268.70	273.90
EITC Amount (*Bakija*)	276.60	1.316	274.00	279.20
EITC Amount (*IRS Paid*)	253.30	1.481	250.41	256.26
Annual Income*	43.43	0.594	42.26	44.60
Annual Earnings*	29.27	0.402	28.48	30.06
Annual Federal tax*	6.68	0.043	6.59	6.76
Less than HS	0.13	0.001	0.13	0.13
High school grad	0.30	0.001	0.29	0.30
Some college	0.28	0.001	0.28	0.28
College degree	0.29	0.001	0.29	0.29
Married	0.30	0.001	0.30	0.30
Single	0.55	0.001	0.55	0.55
White	0.80	0.000	0.80	0.80
Black	0.13	0.000	0.13	0.13
American Indian or Alaska Native	0.01	0.000	0.01	0.01
Asian	0.04	0.000	0.04	0.04
Other	0.02	0.000	0.02	0.02
Hispanic	0.13	0.000	0.13	0.13
Sex (1=male)	0.50	0.001	0.50	0.50
Age	48.11	0.023	48.06	48.16
Number of children <18	0.41	0.002	0.41	0.41
<u>Family means</u>				
EITC Amount (*CPS*)	401.50	2.149	397.30	405.80
EITC Amount (*TAXSIM*)	283.00	1.394	280.20	285.80
EITC Amount (*Bakija*)	288.40	1.389	285.60	291.10

EITC Amount (*IRS Paid*)	347.00	1.749	343.50	350.40
Number of persons	2.21	0.003	2.20	2.21
Number of workers	1.18	0.002	1.18	1.18
Number of qualifying children	0.57	0.002	0.57	0.57
Total income*	61.49	0.191	61.11	61.87
Supplemental Security Income	42.48	0.129	42.22	42.73
Disability Income	179.80	3.528	172.90	186.80
Housing subsidy	8.77	0.178	8.42	9.12
Unemployment compensation	362.80	3.157	356.50	369.00
SNAP benefit	73.66	0.486	72.70	74.62
Public assistance income	47.90	0.931	46.07	49.74

Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015. Shown are means reported using CPS replicate weights that have be recalculated to account for selection into sample. Numbers have been rounded to comply with the Census Bureau's disclosure avoidance guidelines. Dollar amounts are in real 2009 dollars using the personal consumption expenditure deflator.
*In thousands of 2009 dollars

Table 3. Summary Statistics for Linked CPS ASEC-IRS Data, Harmonized Sample of Potential Tax Filers

Full sample	Mean	Standard Err.	95% Confidence Interval	
Reweighted (Obs.= 654.8 thou., N=1,535 mil.)				
EITC Receiver (*CPS*)	0.12	0.001	0.12	0.13
EITC Receiver (*TAXSIM*)	0.13	0.001	0.13	0.13
EITC Receiver (*Bakija*)	0.13	0.001	0.12	0.13
EITC Receiver (*IRS Paid*)	0.15	0.001	0.15	0.15
EITC Receiver (*IRS Eligible)*	0.14	0.001	0.14	0.14
EITC Amount (*CPS*)	244.30	1.580	241.20	247.40
EITC Amount (*TAXSIM*)	245.70	1.586	242.60	248.90
EITC Amount (*Bakija*)	248.40	1.575	245.30	251.50
EITC Amount (*IRS Paid*)	333.20	2.009	329.20	337.20
EITC Amount (*IRS Eligible)*	249.60	1.614	246.40	252.50
Annual Income*	56.19	0.553	55.10	57.30
Annual Earnings*	39.56	0.424	38.72	40.40
Annual Federal tax*	5.98	0.045	5.90	6.07
Less than HS	0.13	0.001	0.13	0.13
High school grad	0.30	0.001	0.30	0.30
Some college	0.28	0.001	0.28	0.28
College degree	0.29	0.001	0.28	0.29
Married	0.40	0.001	0.40	0.40
Single	0.60	0.001	0.60	0.60
White	0.79	0.001	0.79	0.79
Black	0.13	0.001	0.13	0.14
American Indian or Alaska Native	0.01	0.000	0.01	0.01
Asian	0.05	0.000	0.04	0.05
Other	0.02	0.000	0.02	0.02
Hispanic	0.14	0.001	0.14	0.15
Sex (1=male)	0.61	0.001	0.61	0.61
Age	47.80	0.020	47.80	47.90
Number of children <18	0.51	0.002	0.51	0.52

Family means				
EITC Amount (*CPS*)	410.00	2.877	404.00	415.40
EITC Amount (*TAXSIM*)	279.00	1.901	275.30	282.80
EITC Amount (*Bakija*)	280.50	1.883	276.80	284.20
EITC Amount (*IRS Paid*)	422.60	2.736	417.20	428.00
Number of persons	2.37	0.004	2.36	2.38
Number of workers	1.25	0.003	1.25	1.26
Number of qualifying children	0.58	0.002	0.58	0.59
Total income*	63.27	0.225	62.82	63.71
Supplemental Security Income	369.80	4.583	360.80	378.88
Disability Income	209.40	4.660	200.20	218.56
Housing subsidy	9.70	0.208	9.29	10.11
Unemployment compensation	437.10	4.682	427.85	446.35
SNAP benefit	82.36	0.649	81.07	83.64
Public assistance income	53.10	1.207	50.71	55.48

Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015. Shown are means reported using CPS replicate weights that have be recalculated to account for selection into sample. Numbers have been rounded to comply with the Census Bureau’s disclosure avoidance guidelines. Dollar amounts are in real 2009 dollars using the personal consumption expenditure deflator.

*In thousands of 2009 dollars

Table 4. Average survey/admin EITC dollars by respondent type and income quintile

Income Quintile	All 100%	Respondent/ PIK 69%	Respondent/ No PIK 6%	Non-Respondent/ PIK 20%	Non-Respondent/ No PIK 5%	R-P Admin	NR-P Admin
2	224.9	225.9	262	307	76.25	269.2	360.8
3	221.4	186.8	264.7	356.9	53.46	206.1	400.5
4	1028	1054	1115	1022	85.62	880.7	653.5
5	170.2	175.7	246.4	157.1	90.1	297.4	372

Table 5. Comparison of Internal Recipient File to Aggregates in Public-Use IRS Statistics of Income, Tax Year 2006
(numbers in millions)

	(a) Public Use Aggregates		(b) Internal Recipient File		(c) Ratio
Zero Children					
Total benefits	$1,142	2.57%	$1,133	2.63%	0.99
Number of recipients	4.81	20.88%	5.11	22.17%	1.06
One Child					
Total benefits	$16,078	36.22%	$15,750	36.57%	0.98
Number of recipients	8.75	37.98%	8.66	37.57%	0.99
Two Children					
Total benefits	$27,168	61.21%	$26,180	60.80%	0.96
Number of recipients	9.49	41.19%	9.28	40.26%	0.98
Total					
Benefits	$44,388	100%	$43,070	100%	0.97
Recipients	23.05	100%	23.05	100%	1.00

Source: EITC/CP0927 recipient files, 2006, and SOI public reports (U.S. Department of the Treasury, Individual Income Tax Returns 2004). Numbers in column (b) have been rounded to comply with the Census Bureau's disclosure avoidance guidelines.

Table 6. Ratio of Alternative EITC Model Payments to the Internal Recipient File for Tax Year 2006

	Ratio of *CPS* to Internal IRS	Ratio of *TAXSIM* to Internal IRS	Ratio of *Bakija* to Internal IRS	Ratio of *IRS Paid* to Internal IRS	Ratio of *IRS Eligible* to Internal IRS	Ratio of *IRS Paid & Eligible* to Internal IRS
Filing Status						
Head of household						
Total benefits	0.58	0.56	0.57	0.71	0.72	0.65
Number of recipients	0.61	0.61	0.61	0.68	0.73	0.62
Joint						
Total benefits	0.70	1.09	1.36	1.52	1.25	1.06
Number of recipients	0.74	1.19	1.19	1.46	1.19	0.97
Single						
Total benefits	0.50	0.52	0.47	1.41	0.43	0.34
Number of recipients	0.73	0.82	0.61	0.95	0.81	0.47
Qualifying Children						
Zero Children						
Total benefits	2.24	2.58	2.39	6.47	1.92	1.52
Number of recipients	0.95	1.11	0.89	1.32	1.07	0.62
One Child						
Total benefits	0.60	0.63	0.64	0.88	0.83	0.74
Number of recipients	0.58	0.68	0.67	0.82	0.79	0.68
Two Children						
Total benefits	0.53	0.64	0.64	0.79	0.77	0.67
Number of recipients	0.59	0.72	0.72	0.80	0.79	0.67
Total						
Total benefits	0.60	0.68	0.69	0.97	0.82	0.72
Number of recipients	0.67	0.79	0.74	0.92	0.85	0.66

Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2006-2007. Harmonized IRS Eligible sample.

Table 7. Ratio of Alternative EITC Model Payments to the Internal Recipient File for Tax Year 2014

	Ratio of *CPS* to Internal IRS	Ratio of *TAXSIM* to Internal IRS	Ratio of *Bakija* to Internal IRS	Ratio of *IRS Paid* to Internal IRS	Ratio of *IRS Eligible* to Internal IRS	Ratio of *IRS Paid & Eligible* to Internal IRS
Filing Status						
Head of household						
Total benefits	0.54	0.55	0.57	0.74	0.76	0.68
Number of recipients	0.57	0.58	0.58	0.70	0.74	0.64
Joint						
Total benefits	0.72	1.19	1.21	1.39	1.29	1.08
Number of recipients	0.71	1.21	1.27	1.33	1.19	0.96
Single						
Total benefits	0.38	0.41	0.38	0.98	0.35	0.29
Number of recipients	0.60	0.66	0.52	0.78	0.70	0.43
Qualifying Children						
Zero Children						
Total benefits	1.90	2.27	2.22	5.13	1.82	1.50
Number of recipients	0.81	0.96	0.84	1.14	0.97	0.61
One Child						
Total benefits	0.52	0.61	0.63	0.83	0.82	0.72
Number of recipients	0.52	0.62	0.62	0.76	0.76	0.65
Two Children						
Total benefits	0.53	0.68	0.69	0.80	0.82	0.71
Number of recipients	0.56	0.74	0.74	0.79	0.81	0.69
Total						
Total benefits	0.57	0.70	0.71	0.94	0.85	0.74
Number of recipients	0.61	0.75	0.72	0.87	0.83	0.66

Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2014-2015. Harmonized IRS Eligible sample.

Table 8. Effect of EITC on the Official Poverty Rate by Alternative EITC Eligibility Models and Samples

A. Full Sample and CPS ASEC Weight

Year	CPS ASEC	*CPS*	*TAXSIM*	*Bakija*	*IRS Paid*
2005	12.6	11.3	11.3	11.2	11.7
2006	12.3	11.0	11.0	11.0	11.4
2007	12.5	11.0	11.0	11.0	11.5
2008	13.2	12.0	11.9	11.9	12.4
2009	14.3	12.6	12.6	12.6	13.2
2010	15.1	13.4	13.4	13.4	14.0
2011	15.0	13.2	13.2	13.2	13.8
2012	15.0	13.3	13.3	13.3	13.8
2013	14.7	13.0	13.0	13.0	13.6
2014	14.8	13.0	13.0	13.0	13.6

B. Harmonized Sample and Reweighted CPS ASEC Weight for Nonresponse and Link

Year	CPS ASEC	*CPS*	*TAXSIM*	*Bakija*	*IRS Paid*	*IRS Eligible*	*IRS Paid & Eligible*
2005	12.2	11.2	11.3	11.3	11.3	11.4	11.5
2006	12.2	11.1	11.4	11.4	11.3	11.4	11.5
2007	12.1	10.9	11.1	11.1	11.1	11.2	11.3
2008	12.7	11.6	11.8	11.8	11.7	11.9	12.0
2009	14.2	12.7	12.9	12.9	12.9	13.0	13.2
2010	15.2	13.7	14.0	14.0	14.0	14.1	14.3
2011	15.2	13.6	13.9	13.9	13.9	13.9	14.1
2012	15.1	13.6	13.9	13.9	13.8	13.9	14.0
2013	15.0	13.6	13.9	13.9	13.8	13.9	14.0
2014	15.2	13.7	14.0	14.0	14.0	14.0	14.2

Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015

Appendix

The EITC estimation project at the Census Bureau is a joint collaboration with the IRS. Briefly, the goal for IRS has been the estimation of EITC take up versus EITC eligibility. The agency recognized that this estimate requires data that capture household structure, earnings, and income for the full population (when weighted), regardless of filing status. The broad incentive for Census was the availability of what it considered confirmatory data for calibrating its internal tax model and comparing reports of earnings and income to “true” reports. There is some controversy over treating 1040 information as more “truthful” than survey responses; however, for many lower-income tax filers, the parameters that go into EITC eligibility receipt are reported by third parties, making these elements more reliable.[19]

For household structure, both IRS and Census assume the responses from the CPS ASEC, versus dependent claiming in the 1040, as the more likely to reflect the truth. A criticism of the estimation strategy questions that this assumption is valid. The issue of concern is that the CPS ASEC’s questions are asked in March, and household structure may have changed since the preceding tax year. In defense, IRS and Census argue that the complicated nature of credit and dependent qualifications for 1040 filing, leading to the accidental claiming of children outside of the household for EITC, makes the 1040 report more questionable. A child who is claimed for any credit or exemption on a 1040 cannot be claimed on another’s 1040 for a credit or exemption, despite the fact that the rules are different (support versus residency) for each credit. These varying rules lead to confusion among tax filers, who may assume that any child they support—regardless of residency—also qualify them for EITC. In considering which circumstance was more likely (household change versus accidental claiming of non-resident children), IRS and Census researchers determined the latter was the more problematic issue (see Liebman, 2000; McCubbin, 2000; and Holtzblatt and McCubbin, 2003, for the issue of misreporting children for EITC; each paper includes an analysis of IRS audit data).

To further assess the question of households changing structure between the tax year and March, we exploit the longitudinal nature of the CPS ASEC, where approximately half the sample can be linked to their preceding March responses (see Madrian & Lefgren, 1999, and Feng, 2008 for more on linking March samples). Because of the unique identifier we place on the data, we can perform this match either using the PIK or using the probability match as one would with the external-use CPS ASEC.

Table A1 reports how well personal attributes correspond based on the type of match. Sex is used as a matching variable; thus, it is unsurprising that sex is consistent for a matched individual 100 percent of the time for both column 1 and column 3. We also use broad race categories for matching; there is close to a 100 percent chance in all years that exact race also matches between the two years. Hispanic correspondence occurs for 98 to 99 percent. This might be due to a few persons reporting differently in the two different waves, as Hispanic identity seems to be a more fluid characteristic than race and sex (Liebler et al., 2017). In contrast, when looking at PIK matches that don’t match based on probability matching, rates of correspondence range from 99 percent (for race and Hispanic origin) and 83 percent (for sex). The total number of matches constitute about 30 percent of the base-year number of observations.

[19] Examples are W2 wage reports and 1099 reports of investment/dividend income.

Table A1: Two-year CPS ASECs, by PIK match or probability match

		Type of match		
Linked years		(1) Both	(2) PIK only	(3) Prob only
2005-2006	Race	1.000	0.967	0.999
	Hispanic	0.998	0.987	0.995
	Sex	1.000	0.843	1.000
	percent	*89.97*	*1.01*	*9.02*
2006-2007	Race	1.000	0.949	1.000
	Hispanic	0.998	0.988	0.996
	Sex	1.000	0.887	1.000
	percent	*89.63*	*0.76*	*9.61*
2007-2008	Race	1.000	0.976	1.000
	Hispanic	0.999	0.985	0.997
	Sex	1.000	0.836	1.000
	percent	*88.2*	*0.73*	*11.07*
2008-2009	Race	1.000	0.896	1.000
	Hispanic	0.999	0.984	0.994
	Sex	1.000	0.828	1.000
	percent	*88.3*	*0.6*	*11.1*
2009-2010	Race	1.000	0.961	0.999
	Hispanic	0.999	0.984	0.995
	Sex	1.000	0.832	1.000
	percent	*89.35*	*0.61*	*10.04*
2010-2011	Race	1.000	0.933	1.000
	Hispanic	0.999	0.986	0.993
	Sex	1.000	0.870	1.000
	percent	*90.22*	*0.66*	*9.13*
2011-2012	Race	1.000	0.936	0.999
	Hispanic	0.999	0.976	0.995
	Sex	1.000	0.889	1.000
	percent	*89.52*	*0.67*	*9.81*
2012-2013	Race	1.000	0.947	0.997
	Hispanic	0.999	0.971	0.996
	Sex	1.000	0.840	1.000
	percent	*88.6*	*0.71*	*10.68*
20132014	Race	1.000	0.992	0.999
	Hispanic	0.999	0.996	0.995
	Sex	1.000	0.991	1.000
	percent	*78.17*	*10.97*	*10.86*
N		492,200	9,230	56,540
		88.21	1.65	10.13

Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015. Counts are rounded to conform to the U.S. Census Bureau's disclosure avoidance practices.

Table A2. Two-year CPS ASEC samples, observations paid ineligible in year 2 because of income or children out of range

		Ineligibility reason		
Link year		Survey-determined children < claimed	AGI >max	Total
	2005--2006	180	270	460
	2006--2007	210	280	490
	2007--2008	210	260	480
	2008--2009	220	250	470
	2009--2010	220	350	570
	2010--2011	230	230	460
	2011--2012	210	200	410
	2012--2013	190	210	390
	2013--2014	130	130	260
Total		1,790	2,180	3,970

Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015. Counts are rounded to conform to the U.S. Census Bureau's disclosure avoidance practices.

Table A2 reports the number of 1040 filers who are 1) matched across two years of CPS ASEC and 2) are found to be paid and ineligible for EITC in the second year through having a mismatch between their AGI limit and the number of children who determine eligibility (in other words, because the AGI threshold is higher for three eligible children versus two, and two versus one, AGI and children codetermine correct claiming). Generally, we may determine that AGI is too large based on a combination of survey and IRS 1040 income reports.[20] For the approximately four thousand observations who fit the category of an AGI/qualifying child mismatch over the linked samples from 2005-2014, about half claimed more children than appeared to be in their household and about half had too much income than their child claiming permitted.

Focusing on the apparent error in child claiming, between 180 and 230 filers in each second year claim more children for EITC than appear in their household. Of these, the vast majority (70 percent) appear to have zero children in the first year's March survey (Table A3).[21]

[20] Examples: There are spouse earnings (often self-employment) reported for the household but the filer only reports his or her own earnings; occasionally more income from investments or rents are reported in the survey than on the 1040.
[21] Many of these also appear to have zero children in the second year's March survey, but for simplicity the table pools claiming disparities according to the number of children claimed in the second year. For example, those who claim two children in the second year (e.g., tax year 2006)

Approximately 370 persons appeared to have as many or more children in their household in the first year compared with their claiming in the second year (about 21 percent of these filers). The remaining 9 percent have fewer, but non-zero, children in the preceding survey compared with their filing in the second year.

Table A3. Those with survey-determined children < claimed children in table 8, by survey-determined children in year 1

Children claimed	Children in survey, year 1				
Year 2	Zero	One	Two	Three	Total
One	760	210	30	20	1,020
	74.46	*20.73*	*2.85*	*1.96*	*100*
Two	380	140	90	10	620
	61.38	*22.44*	*14.9*	*1.28*	*100*
Three or more	90	20	20	10	150
	62.42	*14.09*	*14.77*	*8.72*	*100*
Total	1,230	370	140	40	1,790
	68.90	*20.77*	*8.04*	*2.29*	*100*
Percent with agreement or +	21%				

Source: combined EITC/CP0927 recipient files, CPS ASEC, Form 1040, and Form W-2, 2005-2015. Counts are rounded to conform to the U.S. Census Bureau's disclosure avoidance practices.

Thus the number that are potentially misclassified as incorrectly paid—approximately 370 filers—constitutes about 5.5 percent of all filers we identify as incorrectly paid EITC in the matched sample (thus about 2 percent of all incorrectly paid filers identified in the pooled CPS ASECs. Of course, this is a lower bound for all filers, since we, by construction, only match a third over each two-year set. On the other hand, we have no way in which to identify changes that occur between March of year one and December of year one—household structure is captured early enough in the tax year that it may not reflect the structure over the full year. An EITC-qualifying child must be resident in the household for more than half the tax year.

In short, especially for the purposes of the current analysis, it is unlikely that misclassifying such a small number of likely paid ineligibles has much of an impact on our estimates. This is especially true for our anti-poverty analysis, since many paid ineligibles share most characteristics with the eligible population, including low income.

References

Feng, S. 2008. "Longitudinal Matching of Recent Current Population Surveys: Methods, Non-Matches and Mismatches." Journal of Economic and Social Measurement, 33(4): 241-252.

Holtzblatt, J. and McCubbin, J., 2003. "Whose child is it anyway? Simplifying the definition of a child." National Tax Journal, 56(3): 701-718.

---

may appear to have zero or one child in their household in the 2007 ASEC to fall under the category "survey-determined children < claimed children."

Liebler, C.A., S.R. Porter, L.E. Fernandez, et al. 2017. “America’s Churning Races: Race and Ethnicity Response Changes Between Census 2000 and the 2010 Census.” Demography, 54(1): 259–284.

Liebman, J. B., 2000. Who Are the Ineligible EITC Recipients. National Tax Journal, 53(4), pp. 1165-1185.

Madrian, B., and L. Lefgren. 1999. “A Note on Longitudinally Matching Current Population Survey (CPS) Respondents.” NBER Working Paper 247.

McCubbin, J., 2000. EITC Noncompliance: The Determinants of the Misreporting of Children. *National Tax Journal*, 53(4, Part 2), pp. 1135-1164