U.S. Census Bureau

EITC Estimates in the CPS ASEC Simulations of After-Tax Income

Hispanic Population

Kathleen Short, Dennis Donahue, George Lynch

SEHSD Working Paper # 2012-19

August 22, 2012



Abstract

The Social, Economic, and Housing Statistics Division of the Census Bureau (SEHSD) regularly releases estimates of household income for the United States using information collected in the Current Population Survey Annual Social and Economic Supplement. As a part of this process, household income after taxes is calculated. The tax calculations use information on family relationships and reported income to simulate tax liabilities and tax credit eligibility for tax units. After-tax income is an important element in the estimation of the Supplemental Poverty Measure. Examination of the effect of the EITC requires information about the actual receipt of the EITC credit, while the tax simulator at the Census Bureau can only imply eligibility. There is evidence to suggest that take-up rates of EITC benefits differ by ethnicity and that among low-income parents Hispanics are less likely to have ever received the tax credit. This analysis uses the 2010 CPS ASEC matched to Tax Year 2009 IRS EITC data. Hispanic tax units are shown to be an important predictor of being a SESHD eligible non-participant.

The views expressed in this research, including those related to statistical, methodological, technical, or operational issues, are solely those of the authors and do not necessarily reflect the official positions or policies of the Census Bureau, or the views of other staff members. The author accepts responsibility for all errors. This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone more limited review than official publications.

The Social, Economic, and Housing Division of the Census Bureau (SEHSD) regularly releases estimates of household income for the United States using information collected in the CPS ASEC. Part of the process involves simulating income taxes. The tax calculations use information on family relationships and reported income to simulate tax liabilities and tax credit eligibility for tax units.¹ After-tax income is an important element in the estimation of the Supplemental Poverty Measure (SPM) described in Short(2011). These estimates provide information about the effect of taxes and transfers on poverty status. In particular, these estimates allow examination of the effect of the Earned Income Tax Credit (EITC) by demographic and socioeconomic characteristics of the population. Examination of the effect of the EITC requires information about the actual receipt of the EITC credit, while the tax simulator at the Census Bureau can only imply eligibility. Insofar as take-up rates differ across demographic groups, or eligibility varies by unobserved characteristics, these estimates may bias our assessment of the implications of policies aimed at alleviating poverty in the U.S.

There is evidence to suggest that take-up rates of EITC benefits differ by ethnicity. Documenting evidence from the 1999 National Survey of America's Families, Phillips (2001) found that low-income Hispanic parents are much less likely to know about the program than low-income non-Hispanic parents of any race. Among low-income parents who know about the EITC, Hispanics are also less likely to have ever received the tax credit.

This paper examines the assignment of EITC benefits by the ASEC tax simulator compared with administrative records of EITC receipt by Hispanic origin. There is evidence that assignments of these benefits are biased upward in the ASEC estimates. The analysis presented here is a first attempt at best answering our questions regarding the participation of Hispanics compared to the total population for the Earned Income Tax Credit. This analysis uses the 2010 CPS ASEC (calendar year 2009) matched to TY 2009 IRS EITC data.²

Table 1 displays aggregate counts and differences of tax-units and credit dollars for the total population and by Hispanic origin.³,⁴ Table 1 shows the number of tax units assigned the EITC by SEHSD who

¹ (O'Hara, 2004 and U.S. Census Bureau, 2004) for details on these calculations.

² The population compared in the following discussion and tables is non-imputed persons on the CPS receiving a Protected Identification Key (PIK) through the Person Identification Validation System (PVS). A PIK is is the mechanism to match CPS ASEC persons to IRS data. All individuals who opt-out of allowing their information to be matched, do not respond to the ACPS ASEC (fl_665 ne "1"), or who cannot otherwise be validated, were dropped. The CPS ASEC weights have been adjusted to account for these dropped records.

³ Missing values for Hispanic were recoded to non-Hispanic.

⁴ The data in this report are from the "Annual Social and Economic Supplement (ASEC)" to the 2010 and 2011 Current Population Survey (CPS). The estimates in this paper (which may be shown in text, figures, and tables) are based on responses from a sample of the population and may differ from actual values because of sampling variability or other factors. As a result, apparent differences between the estimates for two or more groups may not be statistically significant. All comparative statements have undergone statistical testing and are significant at the 90 percent confidence level unless otherwise noted. Standard errors were calculated using replicate weights. Further information about the source and accuracy of the estimates is available at <www.census.gov/hhes/www/p60_238sa.pdf> and <www.census.gov/hhes/www/p60_239sa.pdf>, accessed September 2011.

Table 1: EITC Eligible Tax-Units, Simulated vs. Actual Population: CPS ASEC 2010 Non-imputed, Matchable Tax Units, Weighted				
	TY09			
	CPS Simulation	90 percent C.I. (+/-)	Matched IRS Data	90 percent C.I. (+/-)
Eligible Tax-Units ¹				
Total				
Tax-Units	21,404,000	(179,524)	24,995,000	(216,163)
False Pos. (CPS), False Neg. (IRS)	10,131,000	(132,292)	13,722,000	(160,814)
Hispanic				
Tax-Units	5,683,000	(86,083)	5,999,000	(103,294)
False Pos. (CPS), False Neg. (IRS)	3,016,000	(64,462)	3,332,000	(77,940)
Non-Hispanic				
Tax-Units	15,721,000	(151,140)	18,996,000	(185,721)
False Pos. (CPS), False Neg. (IRS)	7,115,000	(110,771)	10,390,000	(137,300)
Eligible Credit Dollars ²				
Total				
Total	\$42,421,781,000	(\$643,980,810)	\$55,213,901,000	(911,060,370)
False Pos. (CPS), False Neg. (IRS)	\$17,182,611,000	(\$469,944,750)	\$27,959,672,000	(660,220,110)
Absolute Value of Credit Difference ³	\$56,216,272,000	(\$907,776,540)	\$56,216,272,000	(907,776,540)
Hispanic				
Total	\$13,415,444,000	(\$425,220,180)	\$15,351,412,000	(505,852,710)
False Pos. (CPS), False Neg. (IRS)	\$6,583,682,000	(\$294,489,855)	\$7,913,392,000	(377,884,419)
Absolute Value of Credit Difference	\$17,589,422,000	(\$533,870,370)	\$17,589,422,000	(533,870,370)
Non-Hispanic				
Total	\$29,006,337,000	(\$588,075,180)	\$39,862,489,000	(768,777,240)
False Pos. (CPS), False Neg. (IRS)	\$10,598,929,000	(\$384,659,550)	\$20,046,280,000	(519,842,730)
Absolute Value of Credit Difference	\$38,626,851,000	(\$715,084,425)	\$38,626,851,000	(715,084,425)
Notes:				

٦

 $^1\mbox{IRS}$ eligible tax-units are tax units receiving a positive EITC credit from the IRS

²IRS eligible credit dollars are dollars paid to EITC claimaints

³Formula for the absolute value of the difference in credit dollars assigned is: (SUM,i (abs(IRS_i \$ - Sim_i \$)))

Source: TY2009 IRS-CPS ASEC Exact Match

received it according to the IRS records. Overall, 86 percent of tax units who received the EITC were assigned amounts by the ASEC simulator. The figures for Hispanic and non-Hispanic tax units for the two files are also shown. However, these figures mask problems. The analysis separates ASEC and IRS non-overlapping records (tax-units that only receive an EITC credit amount greater than zero in one file) to show the "False Positive" and "False Negative" totals based on IRS data. The percent over-assigned in the ASEC, the non-overlapping unit divided by all assigned tax units, are the false positives. Figure 1 supports the information contained in Table 1. Figure 1 shows the percent with the EITC from both files and the number of units not matching across the files for total and Hispanic tax units⁵.

The first set of columns in the figure show the percent designated as eligible in the ASEC file compared to those who received it in the IRS file. The second set of columns shows the percent assigned the EITC in ASEC who did not have an IRS record of receiving the EITC. The last column set shows consistent participation rates, that is, those assigned EITC by both ASEC and IRS. The figure suggests the possibility that ASEC is over-assigning EITC benefits to Hispanic tax units relative to non-Hispanic units. As such, of course, it is more likely to also correctly assign EITC to those Hispanic units who did receive the benefit.



⁵ Replicate weights are used to calculate the standard errors.

Table 2: EITC Eligible Household Aggregate Credit Amounts

Population: CPS ASEC 2010 Non-imputed, Matchable Households¹, Weighted

Total Population	ТҮ09				
	CPS Simulation	90 percent C.I. (+/-)	Matched IRS Data	90 percent C.I. (+/-)	
Eligible Credit Dollars ²					
Total					
Total	\$42,421,781,120	(521,925,288)	\$55,213,901,360	(750,208,420)	
False Pos. (CPS), False Neg. (IRS)	\$10,431,010,000	(231,905,100)	\$12,004,342,000	(226,408,187)	
Absolute Value of Credit Difference ³	\$45,292,931,280	(608,203,844)	\$45,292,931,280	(608,203,844)	
Hispanic Population					
	CPS Simulation	Matched IRS Data			
Eligible Credit Dollars ²					
Total					
Total	\$13,693,630,170	(354,512,043)	\$15,443,048,910	(421,471,597)	
False Pos. (CPS), False Neg. (IRS)	\$4,571,837,000	(154,319,153)	\$3,158,136,000	(130,748,216)	
Absolute Value of Credit Difference ³	\$14,233,541,060	(376,871,882)	\$14,233,541,060	(376,871,882)	
Non-Hispanic Population					
	CPS Simulation	Matched IRS Data			
Eligible Credit Dollars ²					
Total					
Total	\$28,728,150,960	(487,503,511)	\$39,770,852,450	(625,717,911)	
False Pos. (CPS), False Neg. (IRS)	\$5,859,173,000	(167,844,586)	\$8,846,206,000	(182,581,504)	
Absolute Value of Credit Difference ³	\$31,059,390,220	(495,811,459)	\$31,059,390,220	(495,811,459)	
Notes:					
² IRS eligible rouseroius are nouseroius w	to EITC claimainte		ing an End Creak > at	J	
³ Exercute for the absolute value of the differ	to EIIC Gainants	signed is: (SLIMi	(aba/IRS i & Sim i \$))	١	
Source: TV2000 IPS CPS ASEC Exact M		Signed is. (Solvi,	$(abs(h(O_i \phi - Ohh_i \phi)))$,	
Source. IT 2009 INS-CFS ASEC EXACTIVE	alun				

The bottom half of Table 1 repeats these calculations for dollar amounts assigned. Because the credit amount error is in both directions, it tends to cancel itself out. Therefore, the net difference in aggregate dollars hides the true magnitude of the difference. As a result, to show how the ASEC model compares to IRS data, we use the absolute value of the credit difference. The absolute value of credit difference rows in Table 1 are the sum of the absolute value of the difference in EITC credit amounts assigned by ASEC and IRS. The difference in the amounts is calculated record by record and the absolute values are summed into the aggregate total.

Table 1 showed EITC figures for tax units, but the Census Bureau releases after-tax income estimates for households. Often there are multiple tax units in a household. Table 2 shows EITC eligible households' aggregate statistics. For these calculations an *Eligible Household* is defined as a household that is eligible because at least one tax-unit (person filing taxes or modeled to file taxes) in the household received an

EITC credit. Therefore, an ASEC eligible household flagged at least one person in the household simulated as eligible to receive a positive EITC credit. An IRS EITC household contained at least one taxunit that received an EITC benefit greater than zero⁶.

Table 2 displays aggregate EITC household credit dollar amounts (weighted). Hispanic and non-Hispanic households are analyzed separately. As described above, to show the true magnitude of the difference between the ASEC and IRS data, the absolute value of the difference in credit dollars assigned is reported. Comparing the absolute difference in credit dollars assigned at the tax-unit level to the household level shows that about \$11 billion of the difference in dollars assigned at the person level went to another person in the same household. However, this is only one fifth of the nominal dollars assigned differently. Figure 2 compares the absolute value of credit difference information shown in Table 1 and Table 2.



Table 3 provides estimated participation rates and dollar participation rates of ASEC tax-units by region and ethnicity. Here we consider a person to be eligible if their ASEC EITC credit amount is greater than zero (denominator). A person is a participant if they received a positive EITC benefit in the IRS records (numerator). Therefore, eligible participants in this table are only ASEC eligible persons found to be participants, not all IRS EITC participants. Note that a lower participation rate implies a higher over

⁶ The household EITC credit amount was constructed by (1) Defining the population as non-imputed persons receiving a PIK, (2) Multiplying the tax-unit level ASEC and IRS EITC credit amounts by a weight variable adjusted to account for the removal of imputed persons and persons not receiving a PIK, (3) Summing the weighted EITC credit amounts by household for both ASEC and IRS, (4) Checking to ensure the sum of the weighted household EITC credit dollars equals the sum of the weighted tax-unit level EITC credit dollars in aggregate for both IRS and ASEC

Table 3: EITC P	articipation Ra	ates	abted
	Eligible	,,	Margin
	Tax-units ¹	Rate	(90 +/-) ²
Tax-unit participation rate	21,404,306	53%	1.0%
Region			
New England	754,069	53%	3.8%
East Coast	4,085,138	51%	2.2%
Southeast	3,428,832	55%	2.2%
East Central	3,810,288	58%	2.0%
North Central	1,659,027	58%	2.6%
Southwest	3,777,794	52%	1.9%
West, HI & AI	3,889,159	45%	2.1%
Ethnicity			
Non-Hispanic	15,721,460	55%	1.1%
Hispanic	5,682,846	47%	1.6%
	Eligible Credit		Margin
	Amount ¹	Rate	(90 +/-)2
Tax-Unit Dollar Participation Rate	\$42,421,781,124	49%	1.0%
Region			
New England	\$1,409,755,385	50%	4.2%
East Coast	\$7,917,261,370	49%	2.5%
Southeast	\$6,587,670,942	53%	2.3%
East Central	\$7,238,024,401	54%	2.4%
North Central	\$3,158,309,913	52%	3.5%
Southwest	\$8,361,026,925	48%	2.1%
West, HI & A	\$7,749,732,188	41%	2.3%
Ethnicity			
Non-Hispanic	29,006,337,045	52%	1.2%
Hispanic	13,415,444,078	42%	1.8%
Notes: ¹ IRS eligible tax-units are tax units receivin ² For 90 percent confidence interval.	ng a positive EITC cred	it from the IF	२ऽ

Source: TY2009 IRS-CPS ASEC Exact Match

Г

assignment of EITC benefits by SEHSD. The table shows that participation rates were 53 percent overall and that there were lower rates for Hispanic tax units than for non-Hispanic units, 47 versus 55 percent.

Breakdowns by geographic areas were included in this table to provide context to the participation rates by ethnicity.⁷ As the table shows, participation rates are lower for those residing in the West. This may be true because the West is an area where many Hispanics reside. Additional analysis, however, suggests other factors are important.

Tax-unit dollar participation rates are shown in the bottom section of the Table 3. This shows the percent of assigned dollars in terms of the ASEC tax calculator implying eligibility and the actual amounts paid out by the IRS. For all units, 49 percent of EITC dollars assigned by SEHSD were actually paid out.

Table 4 shows ASEC Eligible Non-Participants by EITC credit amount, AGI, ethnicity, filing status, and filing type. In other words, this group represents those assigned EITC benefits in the ASEC who are not recorded to have received the credit by the IRS data. The credit amount and AGI used are from the ASEC data, not IRS information. The filer and non-filer category are persons with and without 1040s respectively. This table characterizes the over assignment of the EITC by SEHSD in terms of credit dollar amounts and by Adjusted Gross Income (AGI) categories, filing status, ethnicity, and filing unit type as assigned by the ASEC computations. For example, while Hispanic tax units comprised 27 percent of all tax units they made up 30 percent of tax units incorrectly assigned EITC. Further, 21 percent of eligible non-participants were assigned smaller dollar amounts, between \$250 and \$500, while 14 percent had an AGI between \$10,000 and \$12,500.

Since many of the characteristics shown in table 4 overlap, it is useful to use a multivariate approach to examine these eligible non-participating units. Table 5 displays results for a logistic regression that predicts the likelihood that a person is an eligible non-participant. The model builds current intuition and interests into past models (Scholz 1994). These results suggest that the largest predictor of over-assignment by the ASEC simulation is the presence of no qualifying children (the omitted category is one qualifying child). Of interest to us, Hispanic tax units are shown to be an important predictor of being an eligible non-participant. Other characteristics associated with a higher probability of this status are having 3 qualifying children, being married, having a bachelor's degree. Also, note that residing in the West is important, over and above being Hispanic.

⁷ Replicate weights are used to calculate the standard errors.

Summary

This paper provides a preliminary examination of the assignment of EITC benefits to tax units in the CPS ASEC relative to IRS records for tax year 2009. This analysis allows us to characterize elements of the bias in the Census Bureau's ASEC tax simulations. The focus of this analysis was to examine differential bias by ethnicity, particularly focusing on Hispanic units. This was driven by evidence that IRS take up rates for the EITC are lower for this particular group of the population. Examination of other elements of the ASEC tax simulator for this and for other groups would be a useful extension, as well as refinements of the analysis presented here. Our hope is that this and futures analyses will lead to improvements in the assignment of tax liabilities by the Census Bureau SESHD calculations, and that this in turn will lead to an improved understanding of the effect of taxes and transfer payments on low-income households and families in the U.S.

Table 4: ASEC Eligible Non-Participants

CPS ASEC 2010 Non-in	n-imputed, Matchable Persons, Weighted			
	Participants	Margin(+/-) ¹	Total	
TOTAL	10,131,339	(268,593)	-	
Modeled Credit Amount				
<\$100	1,087,362	(88,648)	11%	
<\$250	1,380,349	(97,163)	14%	
<\$500	2,125,897	(137,371)	21%	
<\$1000	615,062	(59,976)	6%	
<\$1500	509,418	(54,005)	5%	
<\$2000	670,560	(54,804)	7%	
<\$2500	643,712	(53,638)	6%	
<\$3000	523,557	(54,230)	5%	
<\$3500	793,777	(69,968)	8%	
<\$4000	285,923	(39,565)	3%	
<\$4500	377,229	(49,330)	4%	
<\$5000	314,011	(37,911)	3%	
\$5000+ ²	804,482	(62,227)	8%	
AGI				
<\$1000	377,445	(55,851)	4%	
<\$3000	643,138	(69,011)	6%	
<\$5000	516,344	(60,334)	5%	
<\$7500	793,536	(67,396)	8%	
<\$10,000	788,219	(78,220)	8%	
<\$12,500	1,433,413	(109,453)	14%	
<\$15,000	606,377	(64,307)	6%	
<\$20,000	955,426	(73,004)	9%	
<\$25,000	840,520	(67,404)	8%	
<\$30,000	813,380	(70,767)	8%	
<\$35,000	807,632	(64,674)	8%	
<\$40,000	860,788	(70,191)	8%	
\$40,000+ ³	680,196	(62,749)	7%	
Filing Status⁴		-		
non-filer	5,944,433	(211,028)	59%	
filer	4,186,906	(156,533)	41%	
Ethnicity		-		
Non-hispanic	7,115,462	(225,531)	70%	
Hispanic	3,015,878	(132,530)	30%	
SEHSD Filing Unit Type ⁴		0		
Joint <65	4,491,720	(149,540)	44%	
Joint, one <65, one >65	267,413	(43,845)	3%	
Head of Household	2,225,385	(113,806)	22%	
Single	3,004,935	(163,707)	30%	
Non-filer	130,089	(29,591)	1%	
Natasi				

¹ For 90 percent confidence interval

² The maximum credit is \$5,657

³ The maximum AGI is \$48,279
⁴ A filer submitted a 1040. A non-filer did not file a 1040 return

⁵ Joint, both >65 not included due to lack of observations

Source: TY2009 IRS-CPS ASEC Exact Match

Table 5: Logistic Regression Results Modeled likelihood of being an eligible non-participant

 $(1 = \text{Eligible Nonparticipant}; 0 = \text{Eligible Participant})^1$ Population: CPS ASEC 2010 Non-imputed, Matchable Persons

	Odds Ratio	90% Wald	
Enect	Point Estimate	Confidence Limits	
Female	0.797	0.756	0.841
Age ³	1.005	1.003	1.007
Hispanic	1.564	1.484	1.647
Race: AIAN only (vs. White)	1.203	1.039	1.393
Race: Asian only (vs. White)	0.921	0.809	1.048
Race: Black only (vs. White)	0.996	0.944	1.051
Race: Other (vs. White)	1.085	0.932	1.262
Foreign Born, Citizen (vs. Native Cit.)	1.028	0.907	1.165
Foreign Born, Non-Citizen (vs. Native Cit.)	1.061	0.954	1.179
West	1.341	1.267	1.420
Married	1.484	1.403	1.569
Bachelor degree or higher (vs. H.S.)	1.473	1.375	1.578
Less than High School (vs. H.S.)	1.228	1.151	1.309
Some College (vs. H.S.)	0.996	0.943	1.051
Claimed 0 Qualifying Children (vs 1)	3.935	3.704	4.181
Claimed 2 Qualifying Children (vs 1)	1.106	1.044	1.171
Claimed 3 Qualifying Children (vs 1)	1.567	1.449	1.696
eit_cred_100 ²	0.996	0.993	0.998
Adjusted Gross Income (logged)	0.876	0.851	0.902
Occupation: Blue Collar vs White Collar ⁴	1.222	1.154	1.294
Occup.: Other Occupation vs White Collar	1.308	1.208	1.417
Public Assistance	1.098	0.952	1.267
Subsidized Lunch	0.736	0.698	0.776
Medicaid	0.753	0.709	0.801
SNAP (food stamps)	0.683	0.648	0.719
	1.010	0.911	1.119
Received Child Tax Credit	1.249	1.179	1.324
Additional Child Tax Credit	0.924	0.860	0.993
State ETIC	1.103	0.667	0.765
	0.714	0.007	0.705
Medicare	0.665	0.582	0.76
	4.070		
exp(intercept)	1.079		
	2343.7		
vvald Pr>x2	<.0001		
Misclassification Rate	0.3049		

Notes:

¹ Eligible if SEHSD EITC credit > 0; Participant if IRS EITC Paid > 0.

² Credit dollar units are calculated in \$100 incrememnts

³ Age is incremented by 5 years in the odds ratio statistic

⁴ Occupation categories aggregated to maintain higher cell counts

Source: TY2009 IRS-CPS ASEC Exact Match

References:

O'Hara, Amy, "New Methods for Simulating CPS Taxes" Amy O'Hara, SEHSD working paper, 2004 <u>http://www.census.gov/hhes/www/income/publications/working.html</u>

Phillips, Katherine Ross, "Who Knows about the Earned Income Tax Credit?, "Urban Institute Series B, No. B-27, January 2001. <u>http://www.urban.org/url.cfm?ID=310035&renderforprint=1</u>

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

Short, Kathleen, *The Research Supplemental Poverty Measure: 2010,* Current Population Reports P60-241, Census Bureau, 2011.

U.S. Census Bureau, "CPS ASEC 2005 Tax Model Documentation", SEHSD Working Paper, 2004. <u>http://www.census.gov/hhes/www/income/publications/working.html</u>