EITC Participation Results and IRS-Census Match Methodology, Tax Year 2021

by

Ciyata Coleman
Internal Revenue Service

Charles Hokayem
U.S. Census Bureau

Sanghun (Eric) Kim
Internal Revenue Service

Ethan Krohn
Internal Revenue Service

Krishnan Patel
Internal Revenue Service

Dean Plueger
Internal Revenue Service

CES 24-75 December 2024

Abstract

The Earned Income Tax Credit (EITC), enacted in 1975, offers a refundable tax credit to low income working families. This paper provides taxpayer and dollar participation estimates for the EITC covering tax year 2021. The estimates derive from an approach that relies on linking the 2022 Current Population Survey Annual Social and Economic Supplement (CPS ASEC) to IRS administrative data. This approach, called the Exact Match, uses survey data to identify EITC eligible taxpayers and IRS administrative data to indicate which eligible taxpayers claimed and received the credit. Overall in tax year 2021 eligible taxpayers participated in the EITC program at a rate of 78 percent while dollar participation was 81 percent.

Keyword: Earned Income Tax Credit, Participation

JEL Classification: H24, H31

[*] Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (Project 6000463: CBDRB-FY24-CES004-016, CBDRB-FY24-CES026-0014, CBDRB-FY24-CES004-018). The views expressed in this paper are those of the authors and do not necessarily represent the views of the IRS.

Summary of Earned Income Tax Credit Participation for Tax Year 2021

The American Rescue Plan Act (ARPA) modified tax year 2021 Earned Income Tax Credit (EITC) eligibility rules by expanding the eligible age range for working taxpayers with no qualifying children (often referred as the "childless EITC").[1] ARPA also increased the amount of the credit these taxpayers were eligible to receive.[2] The tax law changes were effective only for tax year 2021. Because the eligibility rules for tax year 2021 differ from other tax years, information for tax years 2019 and 2020 are also included in this paper.

Table 1 presents taxpayer and dollar participation rates for tax years 2019 to 2021. Eligible taxpayers participated in the EITC program at a rate of 78 percent in tax year 2021. [3] The dollar participation rate was 81 percent. The taxpayer participation rate was 82 percent in tax year 2019 and 80 percent in tax year 2020 while the dollar participation rate was 88 percent in tax year 2019 and 85 percent in tax year 2020. The number of taxpayers eligible for the credit in tax year 2021 increased substantially from prior years (about 26 million 2021 from about 20 million in tax years 2019 and 2020).[4]

Table 1: Total Tax Units Modeled Eligible, Modeled Dollar Amounts, and Participation Rates, Tax Years 2019-2021

Tax Year	Tax Units Modeled Eligible	Taxpayer Participation Rate	EITC Dollars Modeled Eligible[5]	Dollar Participation Rate
2019	19.8 million	82 percent	$37.5 billion	88 percent
2020	20.1 million	80 percent	$38.4 billion	85 percent
2021	25.9 million	78 percent	$42.9 billion	81 percent

Source: Tax Years 2019-2021 IRS-CPS Exact Match. DRB Approval Numbers: CBDRB-FY24-CES004-016, CBDRB-FY24-CES026-0014, CBDRB-FY24-CES004-018.

[1] In tax years other than tax year 2021, the eligible age range for childless workers is 25-64. For tax year 2021, taxpayers over age 65 and taxpayers between ages 19-25 were also eligible so long as they met other EITC eligibility criteria. Taxpayers who were 18 years old were also eligible if they were homeless.

[2] In tax year 2021, the maximum EITC amount for childless workers was $1,502. In tax year 2020, the maximum amount was $538.

[3] Definitions of taxpayer and dollar participation rates as well as methods used to derive participation rates can be found in the Methodology Section at the end of the paper. This paper uses the terms "taxpayers" and "tax units" interchangeably. The estimates presented in this paper reflect a refinement to the EITC methodology that excludes filers with a "B" return processing code. More information about this refinement can be found in the Methodology Section.

[4] The number of taxpayers actually paid EITC also increased over this period. In tax year 2019 and 2022 about 26 million taxpayers were allowed EITC. In tax year 2021, the number of taxpayers paid EITC increased to 32 million.

[5] It is possible the estimated eligible dollar amounts in this table are understated as the current weighting procedure tends to underestimate the actual amount of EITC paid to taxpayers. For example, in tax year 2021 the understatement was about 9 percent ($59 B modeled paid versus $65 B actually paid).

Socioeconomic factors important to EITC participation span a broad range of categories. By filing status females with or without dependents participate at higher rates than single males, and females with dependents participate at rates well above married taxpayers with dependents.[6] Married taxpayers participate more than single taxpayers and the difference is statistically significant (81 percent and 78 percent, respectively). The difference in participation rates for persons with three or more qualifying children (85 percent) is not statistically significant from those with one or two qualifying children (87 percent and 87 percent, respectively).

In tax year 2021 tax units with adjusted gross incomes below $12,500 and modeled credit amounts under $1,000 comprise about 54 percent and 56 percent, respectively, of eligible non-claimants. Of the eligible non-claimants, about 29 percent had children, and 19 percent filed a return.

Table 2 shows the number of eligible non-claimants by filing status for tax years 2019 to 2021.[7] In tax year 2021, 5.7 M, or about 22 percent of eligible tax units did not participate. In tax year 2019 and tax year 2020 the number of eligible non-claimants was estimated to be 3.5 M and 4.1 M, respectively.[8] The increase in the number of non-claimants over the period was driven by an increase in the number of eligible non-filers. In tax year 2019, the number of eligible non-filers was estimated to be 2.3 M. By tax year 2020, the number had increased to 3.2 M, and by tax year 2021 the number was 4.6 M. The number of non-claimant filers ranged between 0.9 M and 1.2 M over this period.

[6] The participation rate for females with and without dependents is 90 percent and 75 percent, respectively. The participation rate for single males is 69 percent. Married persons with dependents participate at a rate of 84 percent.

[7] For an alternative estimate of eligible non-claimants using a different approach relying only on IRS administrative data, refer to IRS Working Paper "Tax Year 2022 EITC Benefits Gap Estimate for Filers Using IRS Administrative Data.".

[8] The increase in the number of non-claimants for tax year 2021 relative to prior years is due to the temporary expansion of the EITC under ARPA for workers with no qualifying children. There was a 7 M (100 percent) increase in the number of eligible taxpayers with no qualifying children between tax year 2020 and tax year 2021. Taxpayers eligible for childless EITC participate at lower rates than taxpayers with qualifying children. For example, in tax year 2020 and 2021, taxpayers with no qualifying children saw participation rates of 68 percent and 71 percent, respectively. Taxpayers with qualifying children had an 86 percent participation rate in both years. In TY2022, the number of non-claimants will likely return to volumes seen in tax year 2019 and tax year 2020.

Table 2: Eligible Non-Claimants, Tax Years 2019-2021

Tax Year	Non-Filers	Filers	Total
2019	2.3 million	1.2 million	3.5 million
2020	3.2 million	0.9 million	4.1 million
2021	4.6 million	1.1 million	5.7 million

Source: Tax Years 2019-2021 IRS-CPS Exact Match. DRB Approval Numbers: CBDRB-FY24-CES004-016, CBDRB-FY24-CES026-0014, CBDRB-FY24-CES004-018.

Table 3 presents the unclaimed EITC dollars for tax years 2019 to 2021.[9] The first two columns show unclaimed EITC dollars for non-filers and filers who claimed no EITC, respectively. The third column shows unclaimed EITC dollars for under-claimants.[10] In tax year 2021, unclaimed EITC dollars is estimated to be $8.2 B which includes $5.6 B in unclaimed dollars for non-filers, $1.3 B in unclaimed dollars for filers who claimed no EITC, and $1.3 B in unclaimed dollars for under-claimants. In tax year 2020, unclaimed EITC dollars is estimated to be $5.9 B which includes $3.7 B in unclaimed dollars for non-filers, $1.2 B in unclaimed dollars for filers who claimed no EITC, and $1.0 B in unclaimed dollars for under-claimants. In tax year 2019, unclaimed EITC dollars is estimated to be $4.7 B which includes $3.0 B in unclaimed dollars for non-filers, $0.9 B in unclaimed dollars for filers who claimed no EITC, and $0.8 B in unclaimed dollars for under-claimants.[11]

Table 3: Unclaimed EITC Dollars, Tax Years 2019-2021

Tax Year	Non-Filers	Filers Who Claimed No EITC	Under-Claimants	Total
2019	$3.0 billion	$0.9 billion	$0.8 billion	$4.7 billion
2020	$3.7 billion	$1.2 billion	$1.0 billion	$5.9 billion
2021	$5.6 billion	$1.3 billion	$1.3 billion	$8.2 billion

Source: Tax Years 2019-2021 IRS-CPS Exact Match. DRB Approval Numbers: CBDRB-FY24-CES004-016, CBDRB-FY24-CES026-0014, CBDRB-FY24-CES004-018.

[9] For an alternative estimate of unclaimed EITC dollars relying only on IRS administrative data, refer to IRS Working Paper "Tax Year 2022 EITC Benefits Gap Estimate for Filers Using IRS Administrative Data."
[10] Under-claiming can occur when the actual EITC amount paid is less than the modeled EITC amount which can come from differing estimates of income and/or number of qualifying children.
[11] The increase in unclaimed EITC dollars for tax year 2021 relative to prior years is due to the temporary expansion of the EITC under ARPA for workers with no qualifying children.

The following section provides more information about the methodology used to derive the estimates.

Methodology Overview

The EITC methodology provides two primary statistics to gauge EITC participation: (1) the taxpayer participation rate and (2) the dollar participation rate. The taxpayer participation rate (a percentage) is the total number of taxpayers who received an EITC payment divided by the total number of eligible taxpayers. The dollar participation rate (a percentage) is the total amount of eligible EITC dollars paid out divided by the total EITC dollars for which all taxpayers are eligible. Both participation rates represent participation conditional on eligibility.[12]

The Exact Match methodology for estimating the participation rates consists of linking individuals in the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) to their tax return information from IRS administrative data. The CPS ASEC identifies the EITC eligible taxpayers and eligible dollars, while the IRS administrative data and return information indicate which taxpayers claimed and received the credit and dollars paid. The following sections describe the CPS ASEC and IRS administrative data and detail the process of producing the participation rates.

Current Population Survey Annual Social and Economic Supplement

The CPS ASEC is a stratified random sample of 60,000-90,000 household addresses from the civilian, non-institutionalized population in the United States and is collected by the Census Bureau. It is a supplement to the monthly Current Population Survey, conducted in February-April of each year, that contains detailed information on demographic characteristics, annual earnings and incomes, employment, hours worked, and family structure. It also serves as the official source of U.S. income and poverty statistics. All income information in the CPS ASEC refers to the prior calendar year while demographic information is as of the survey interview date. Since the CPS ASEC does not collect tax information, the IRS provided a data file based on the public use CPS ASEC that contains modeled tax returns, including adjusted gross income, earned income, number of qualifying children, the EITC amount, and other important tax return fields.

[12] Taxpayers who receive the credit but are not modeled eligible are not included in the participation rates.

IRS Administrative Data

The IRS provided the Census Bureau administrative tax data and return information. The return information covers all returns through cycle 26 of the year after returns are filed (e.g., tax year 2021 returns posted through cycle 26 of 2023), allowing for many audits to be completed and late filed tax returns with EITC to be processed. The key variables include number of EITC qualifying children, claimed EITC amount, and final EITC paid amount. Other important variables include wages, tips, and other compensation before any payroll deductions from all jobs in each tax year, adjusted gross income, investment income, and self-employment income.

Modeling EITC Eligibility and Amount in the CPS ASEC

The first step of the participation methodology involves modeling tax units, EITC eligibility, and EITC credit amount using the CPS ASEC. Since the CPS ASEC does not collect any information on tax filing behavior, the IRS applies internal programming to the CPS ASEC to produce estimates of eligibility. This internal programming is the IRS tax model.[13] This section outlines the process of creating tax units and modeling EITC eligibility.

Analysts in the Taxpayer Services, Strategies and Solutions program area at IRS download the public use CPS ASEC data file that is associated with the relevant tax year. For example, the 2022 CPS ASEC file captures income sources related to calendar year 2021 and is used to generate TY2021 estimates of eligibility. After downloading the files, analysts apply programming to the file to derive tax units and estimates of EITC eligibility. The programming is updated annually to capture inflation adjustments and relevant changes to the tax law.

To create tax units, the CPS ASEC person file is first divided into married and non-married individuals. Records for married individuals are merged to create a combined record for both persons. Non-married individuals are categorized into four groups: (1) children under 19; (2) adults from 19 to 23 who are full-time students; (3) disabled adults living with a parent or sibling; and (4) all other adults. Group 4 is combined with married tax units to create the set of tax units.[14] Groups 1-3 are used to determine the initial number of qualifying children for each tax unit.

[13] The participation methodology in this paper does not use the EITC eligibility estimates created by the CPS ASEC tax model maintained by the Census Bureau. National estimates of eligibility from the two models are similar. For example, in TY2021, the IRS model estimated 25.8 M eligible tax units, and the Census CPS ASEC tax model estimated 27.3 M eligible.

[14] Between TY2018 and TY2021, this process generated about 175 M tax units per year, ranging from 173 M to 177 M.

After deriving the tax units, the income tax concepts needed to model eligibly are created. For example, earned income, investment income, and adjusted gross income (AGI) of the tax unit are estimated by combining the various income sources captured by the survey.[15] Additionally, other tax credit amounts such as Child and Dependent Care Credit (CDCC), American Opportunity Tax Credit (AOTC), refundable and non-refundable portions of the Child Tax Credit (CTC), and the Credit for Other Dependents (ODC) are estimated by applying the various rules of eligibility for these credits.

Persons in households with multi-generational families may allocate qualifying children amongst legally entitled tax units.[16] The model identifies households for qualifying child reallocation when the parental tax unit receives no additional tax benefit by claiming the qualifying child, and there are other tax units within the same household who would benefit from filing with additional qualifying children.[17] Tax units who are ineligible for EITC under all circumstances (for example, AGI or earned income is too high) or tax units whose tax liability would increase (or receive a smaller refund) due to the reallocation of qualifying children, are not considered in this process. The model also reallocates qualifying children between cohabitating parents.[18]

Finally, the programming estimates EITC amounts for tax units modeled as eligible based on their estimated AGI, earned income, marital status, and the number of qualifying children after reallocation. The IRS provides the Census Bureau with a final file containing these variables to be linked with administrative tax data.

Linking CPS ASEC to Administrative Tax Data

The Economic Reimbursable Surveys Division (ERD) at the Census Bureau processes the CPS ASEC and administrative tax data from IRS to add a Protected Identification Key (PIK) that allows the two data sources to be linked.[19] ERD assigns the PIK using a proprietary linkage software, the Person Identification Validation System (PVS). PVS assigns the PIK in the administrative tax data using Social Security Number. Since the Social Security Number is not collected in the CPS ASEC, PVS assigns the PIK using a

[15] For example, the AGI estimate is derived by summing the wages, self-employment income, disability income, taxable Social Security, and investments, etc. for both the primary and secondary filer, if present.

[16] See example 1, page 13 of IRS Publication 596 (TY2023) <https://core.publish.no.irs.gov/pubs/pdf/p596--2023-00-00.pdf>

[17] For example, consider a household consisting of a 50-year-old woman who earns $15,000 annually (who is the grandmother of the child living in the household), and an unemployed (no income) 30-year-old mother with a 5-year-old daughter. The 30-year-old mother receives no benefit from filing a return and claiming her child for EITC; however, the grandmother receives EITC from filing a return and claiming the grandchild as a qualifying child. In this instance, the household is better off when the grandmother claims the child for EITC.

[18] Between TY2019 and TY2021, 4-8% of the population modeled eligible for EITC had a qualifying child reallocated.

[19] The PIK acts as a unique identifier for individuals in the CPS ASEC and administrative tax data.

probabilistic matching model based on name, address, date of birth, and gender.[20] The PIK assignment rate for tax year 2021 is 83.7 percent. To make the resulting linked CPS ASEC sample representative of the U.S. population, the weights of the linked individuals are adjusted, or re-weighted, to allow for PIK non-assignment (re-weighting is described later).

Adjustments to Original Modeled Eligibility

While the IRS programming applied to the CPS ASEC described above (i.e., IRS tax model) provides initial estimates of credit eligibility and amount, these initial estimates are further refined using IRS administrative data to account for tax behavior not predicted by the model and to correct for underreporting of income in the CPS ASEC. The following series of non-income and income adjustments are implemented in a cumulative manner:

1. **Change filer from head to spouse adjustment:** The IRS tax model designates the head of a household as the primary filer, which may not always reflect actual filing behavior. This adjustment moves the modeled tax return information from the head to the spouse if the spouse filed but the head did not. If both the head and the spouse filed, the modeled return information is moved from the head to the spouse when the spouse claimed EITC, and the head did not. If both the head and spouse filed, and both claimed EITC, the modeled return information is moved from the head to the spouse when the spouse claimed the qualifying child and the head did not. If both filed and neither claimed EITC, but a qualifying child appears on the spouse's return, the modeled return information is moved to the spouse.

2. **Drop "married filing separate" and sanctioned adjustments:** These adjustments set the modeled credit amount to zero for couples where the modeled filer actually filed as "married filing separately" and for situations where the modeled filer submitted a claim but was denied payment.[21]

3. **Qualifying child (QC) adjustments for cohabiting partners, multi-generational families, and all others:** This adjustment alters the number of qualifying children assigned to tax units and modeled filers by comparing the number of qualifying children reported on the survey to the

[20] For more information about the PVS system refer to 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.

[21] In tax year 2021, a change in the tax law allowed some "married filing separate" filers to claim EITC in that year. Since this change only affects a small number of filers, the methodology still drops this group.

number reported on the tax return. Important to this process is the creation of a "QC unit" within a household. A "QC unit" contains all individuals in the household who could share a qualifying child. We make the simplifying assumption that any child identified by the model as a qualifying child could be claimed by any relative, meaning a "QC unit" consists of all related individuals in the household. We also consider non-married partners (and their children) as part of the householder's "QC unit" (i.e., they are treated as relatives of the householder).

We adjust "QC units" based on the presence of "surplus" and "deficit" children. When a filer is modeled to have more qualifying children than they actually claimed on their return (i.e., more children in CPS ASEC than claimed on return), they have "surplus" children. When a filer claimed more children than modeled, they have "deficit" children. Non-filers are considered to not have claimed any children, so are classified as having a "surplus" if they are modeled to have any qualifying children. Adjustments are only made to units with at least one "surplus" qualifying child—units where, according to the CPS ASEC, there is at least one more qualifying child present than was actually claimed on the return. Surplus and deficit units are handled as follows based on filing, surplus, and deficit status:

- **Filer with a surplus:** The number of qualifying children is reduced to the number of qualifying children claimed on the return. The filer's eligibility and credit amount are re-estimated based on the reduced number of qualifying children.
- **Non-filer with a surplus:** The number of qualifying children is reduced by the amount necessary to meet the needs of any deficit filers in the household (i.e., units in the household that claimed more qualifying children than modeled for them). In households comprised of both filers and non-filers with surplus children, deficits are first met by taking surplus children from filers followed by non-filers. As an example, if a household has a filer with 1 surplus child, a non-filer with 2 surplus children, and a filer with 1 deficit child, the child needed to make up the deficit comes from the filer with a surplus child (reducing the number of qualifying children for that filer by 1) rather than the non-filer. If the filer had a deficit of 2 children (rather than 1 deficit child), the filer with 1 surplus child still has that number reduced by 1, and the non-filer will have their number reduced by 1. If qualifying children change during this process, eligibility and credit amount are re-estimated based on the new number of qualifying children.
- **Filer with a deficit:** The number of qualifying children is increased to the number of children claimed on the return. If the number of surplus children in the unit is less than

the number of deficit children, deficits are met only to the extent a surplus allows. As an example, if a household has a filer with 2 deficit children and another filer with only 1 surplus child, the deficit filer's number of qualifying children is increased by 1. If qualifying children change during this process, eligibility and credit amount are re-estimated based on the new number of qualifying children.

The qualifying child adjustments described above apply to all "QC unit" groups which include cohabiting partners, multi-generational families, and all others.

4. **Income adjustment:** This adjustment helps correct for potential income under-reporting in the CPS ASEC and uses the best measure of wages to estimate EITC eligibility. All CPS ASEC wages are replaced with W-2 wages. Single individuals with no W-2 wages receive a wage value of 0 unless they filed a 1040 return, in which case they receive the amount of wage and salary income from the 1040 return. For married couples, 1040 wage and salary income is used only if W-2 wages are not found for either spouse. We use any modified wages to adjust modeled earnings and modeled adjusted gross income.
We also perform an additional income modification for investment income. If the investment income on the 1040 return exceeds the modeled investment income, the modeled income is replaced with the 1040 investment amount and modeled adjusted gross income is recomputed to account for the higher investment amount. In a similar way, if the non-wage non-investment portion of adjusted gross income is higher on the 1040 return than the corresponding portion of modeled adjusted gross income, the modeled adjusted gross income is recomputed to account for the higher 1040 amount. Following these income adjustments, eligibility and credit amount are re-estimated with the new wage and income amounts.

5. **Add childless filer adjustment:** The IRS tax model treats all married persons and most single adults as potential filers. If a single adult is not modeled as a potential filer, but actually did file, they are included in the population of potential filers. The creation of new potential filers begins during the various "QC" adjustments, where single adult filers who are not modeled as potential filers *but claim children* are treated as "deficit" filers (i.e., their number of modeled children, zero, is less than their number of claimed children). If their deficit can be reduced by taking "surplus" children, they become new potential filers. This adjustment only applies to single adult filers not modeled as potential filers *who do not claim children*. They are converted to new

potential filers with zero qualifying children. Following this adjustment, the new potential filers' eligibility and credit amount are estimated.

6. **Drop "B" filers:** This adjustment, which was first introduced when estimating the participation rate for tax year 2021, drops filers with a "B" return processing code which indicates the filer "does not want or qualify for EITC." The "B" code can be generated if a filer excludes themselves from receiving EITC for religious reasons (e.g., Amish or Mennonites), has only unearned income (e.g., child support), or may be the qualifying EITC child of someone else. The "B" code can also be generated by tax preparation software if the software determines the filer is not eligible based on information provided during the completion of the return.

   Applying this new adjustment to tax years 2019 and 2020 leads to revised participation estimates for these years. With the revision, the taxpayer participation rate is 82 percent in tax year 2019 and 80 percent in tax year 2020 while the dollar participation rate is 88 percent in tax year 2019 and 85 percent in tax year 2020. The addition of the "B" adjustment caused the taxpayer participation rate to increase between 4 and 5 percentage points during tax years 2019 to 2021, while the change caused the dollar participation rate to increase between 3 and 4 percentage points over the same time period.

Weighting Procedure

The CPS ASEC includes supplement weights designed to sum to the total civilian, non-institutionalized population in the United States. Individuals who are not assigned a PIK are removed from the analysis.[22] Removing these individuals from the analysis potentially causes the sample to no longer be representative, and the supplement weights will not sum to the total population. To address these concerns, we assume PIK non-assignment is missing at random conditional on race, ethnicity, age, and marital status and inflate the supplement weights for the individuals that are assigned a PIK. The factors used to inflate the supplement weight come from dividing the sample into groups defined by race, ethnicity, age, and marital status. After calculating factors for each group, we apply the factor to each supplement weight, creating a new set of adjusted weights used throughout the analysis.

[22] Refer to earlier section "Linking CPS ASEC to Administrative Tax Data" for a description of the PIK assignment process.